

**“Maybe We Can Just You Know See How It’s Relevant” –
The Use of *You Know* as a Discourse Marker in
Academic ELF Interaction**

Heini-Marja Rintaniemi
University of Tampere
Faculty of Communication Sciences
Master’s Programme in English Language and Literature
MA Thesis
May 2017

Tampereen yliopisto
Viestintätieteiden tiedekunta
Englannin kielen ja kirjallisuuden maisteriohjelma

RINTANIEMI, HEINI-MARJA: "Maybe We Can Just You Know See How It's Relevant" – The Use of *You Know* as a Discourse Marker in Academic ELF Interaction

Pro gradu -tutkielma, 94 sivua + liite [1 sivu]

Toukokuu 2017

Tässä pro gradu -tutkielmassa tarkastellaan, kuinka lingua franca -englantia (ELF:ää) puhuvat käyttävät diskurssipartikkeli *you know*'ta akateemisissa diskurssissa. Diskurssipartikkeleita, kuten *you know*'ta on tutkittu paljon englantia äidinkielenään puhuvien keskuudessa, mutta hiljalleen englantia vieraana kielenä puhuvat ovat myös nousseet kiinnostuksen kohteeksi. Usein tutkimuskohteena ovat kuitenkin olleet englannin kielen oppijat eivätkä niinkään englannin jo hyvin hallitsevat kompetentit puhujat. Niinpä tämän tutkielman tarkoituksena on tuoda *you know*'n tutkimukseen toisenlaista näkökulmaa selvittämällä, miten sujuvasti englantia vieraana kielenä puhuvat käyttävät diskurssipartikkeli *you know*'ta ja miten saadut tutkimustulokset mahdollisesti eroavat aiemmista tutkimustuloksista.

Tutkimusaineistona on käytetty ELFA-korpusta, joka koostuu akateemisesta lingua franca -englannista (ELFA 2008). ELFA (2008) on karkeasti litteroitu miljoonan sanan korpus, jonka data on nauhoitettu suomalaisissa yliopistoissa. Korpuksessa on n. 650 puhujaa, joilla on 51 eri äidinkieltä (ELFA 2008). Korpuksen akateemisen luonteen vuoksi oletuksena on, että jokainen ELF-puhuja hallitsee englannin kielen hyvin, koska he voivat käydä akateemista keskustelua kyseisen kielen välityksellä.

You know'ta lähestytään tässä tutkielmassa kahdella eri tavalla. Ensiksi tutkielmassa selvitetään, miten ELF-puhujat käyttävät *you know*'ta eli mitä funktioita sillä on ELFA-korpuksessa. Sen jälkeen saatuja tuloksia verrataan erityisesti kahteen aiempaan tutkimukseen, Mülleriin (2005) sekä Houseen (2009), joissa tutkimuskohteena ovat niin ikään englantia vieraana kielenä puhuvat, joilla on akateeminen tausta. Tutkimustulokset tukevat pääpiirteittäin Müllerin (2005) ja Housen (2009) tutkimuksia, mutta selkeitä eroja on havaittavissa. House (2009) mm. toteaa, etteivät ELF-puhujat käytä *you know*'ta lainkaan vuorovaikutussuhteiden luomiseen kuulijan kanssa, vaan he käyttävät sitä pääasiallisesti tekstitasolla auttaakseen itseään sujuvamman puheen tuottamisessa. Müller (2005) puolestaan toteaa *you know*'lla olevan tasapuolisesti sekä vuorovaikutuksellisia että tekstuaalisia piirteitä. Tämän tutkielman tulokset tukevat Housea (2009) siinä mielessä, että valtaosa ELFA-korpuksen *you know*'ista operoi tekstitasolla. Toisaalta jäljelle jäänyt osa kuitenkin osoittaa, että *you know*'lla on tekstuaalisten piirteiden lisäksi myös vuorovaikutuksellisia piirteitä ELF-puheessa (vrt. Müller 2005). Näiden eroavaisuuksien vuoksi on selkeää, että *you know*'n funktioita on syytä tutkia entisestään myös kompetenttien ELF-puhujien keskuudessa.

Avainsanat: you know, diskurssipartikkeli, ELF, ELFA

Table of Contents

| | | |
|--------|---|----|
| 1. | Introduction..... | 1 |
| 2. | Theoretical background..... | 5 |
| 2.1 | Relevant terminology | 5 |
| 2.2 | Defining the functions of <i>you know</i> | 7 |
| 2.2.1 | Marking lexical or content search..... | 11 |
| 2.2.2 | Marking repair..... | 13 |
| 2.2.3 | Introducing an explanation..... | 15 |
| 2.2.4 | Quotative <i>you know</i> | 18 |
| 2.2.5 | Linking propositions together | 20 |
| 2.2.6 | Highlighting certain elements | 21 |
| 2.2.7 | Securing comprehension | 23 |
| 2.2.8 | Appeal for understanding..... | 26 |
| 2.2.9 | Reference to shared knowledge | 27 |
| 2.2.10 | Acknowledge that the speaker is right | 29 |
| 3. | Corpus data and methodology..... | 30 |
| 3.1 | ELFA | 31 |
| 3.2 | Method of analysis | 32 |
| 4. | How is <i>you know</i> used by academic ELF speakers? | 34 |
| 4.1 | Functions of <i>you know</i> in the ELFA corpus | 36 |
| 4.1.1 | Marking lexical or content search..... | 38 |
| 4.1.2 | Marking repair..... | 44 |

| | | |
|--------|---|----|
| 4.1.3 | Introducing an explanation..... | 47 |
| 4.1.4 | Quotative <i>you know</i> | 50 |
| 4.1.5 | Linking propositions together | 52 |
| 4.1.6 | Highlighting certain elements | 55 |
| 4.1.7 | Securing comprehension | 62 |
| 4.1.8 | Appeal for understanding..... | 67 |
| 4.1.9 | Reference to shared knowledge | 70 |
| 4.1.10 | Acknowledge that the speaker is right | 73 |
| 4.1.11 | Unclassified instances | 76 |
| 4.2 | Comparison to previous literature | 77 |
| 5. | Conclusions..... | 88 |
| | Works Cited | 92 |
| | Appendix 1. Abbreviations used in the transcript of the ELFA corpus | 95 |

1. Introduction

Discourse markers are a distinctive feature of spoken language, yet even today some regard them as unnecessary fillers, markers of poorly planned speech that should be avoided. Many previous studies (e.g. Östman 1981; Schiffrin 1987; Jucker & Ziv 1998), however, have shown that discourse markers have a relevant function in spoken language – and it is not to fill in a gap because the speaker does not have anything better to say. In fact, as this thesis will also show, it is quite the opposite.

Discourse markers have been a study of interest from the 1970s onwards (Schourup 1999: 229). Since then they have been studied from various perspectives, but their past has been quite complex (see, e.g., Jucker & Ziv 1998: 1–2). The term itself is a disputed one: they have been called *discourse markers* (e.g. Schiffrin 1987), *pragmatic particles* (e.g. Östman 1981), *pragmatic expressions* (e.g. Erman 1987) or *pragmatic markers* (e.g. Erman 2001), just to name a few. Furthermore, the different functions discourse markers may have are also vast. These functions include, e.g., the aforementioned *filler* as well as *discourse connector*, *confirmation-seeker*, *hesitation marker* and *repair marker* (Jucker & Ziv 1998: 1). However, despite the difficulties in defining the term and its functions, discourse markers persistently hold their ground as an interesting and relevant topic among present studies. For example, Gánem-Gutiérrez and Roehr (2011) study how Spanish learners of English use discourse markers, and Fuller (2003) examines how the use of *you know* is subject to change depending on the roles the speaker takes during interviews and casual conversation.

The discourse marker this thesis is focusing on is *you know*. Like the term *discourse marker*, *you know* has also been given many different names based on its functions, such as *gambit*, *pragmatic particle*, *discourse operator*, *relational marker* and *discourse marker* (see, e.g., House 2009: 172). The term *discourse marker* is utilised in this thesis simply because, as Schourup (1999: 228) points out, it is the most commonly used term in the field of research. In this thesis, discourse markers are defined as linguistic elements that, e.g., have little or no semantic meaning and are syntactically

optional (Müller 2005: 27). Therefore, the term *discourse marker* excludes utterances such as *You know me too well* where *you know* functions in its original, lexical sense.

Like discourse markers in general, the discourse marker *you know* is not an uncommon subject of interest as it has been studied from the 1980s onwards (e.g. Östman 1981; Erman 1987; Schiffrin 1987). Until recently, the vast majority of studies have concentrated on native speakers of English and their use of *you know* as a discourse marker (e.g. Östman 1981; Schiffrin 1987; Erman 1987, 2001; Redeker 1990; Jucker & Smith 1998). Currently, however, the focus seems to have shifted to study how *you know* is used by non-native English speakers, learners in particular (e.g. Fung & Carter 2007; Polat 2011). Few researcher, though, have shown an interest in studying how *you know* is used as a discourse marker by *fluent* non-native speakers of English. Hence, there is a need in current research to study the discourse marker *you know* from a different perspective, which is exactly the aim of this thesis: to study how fluent non-native English speakers utilise the discourse marker *you know* in interaction.

However, as mentioned above, some researchers have studied *you know* in non-native English interaction, two of which are particularly important for this study, Müller (2005) and House (2009). Müller (2005) studies four different discourse markers, *you know* being one of them, and establishes ten different functions for *you know*, half working on textual level and the other half on interactional level. When *you know* functions on textual level, it is primarily a speaker-oriented device, meaning that *you know* is not directed to the addressee. When *you know* functions on interactional level, however, its function is primarily to elicit cooperation from the addressee (Müller 2005: 30–31.) House (2009), on the other hand, argues that although *you know* is considered primarily as an interactional, or as she calls it interpersonal, device in native English speakers' use, it is not the case in non-native English interaction. In fact, House (2009) emphasises that the main function of *you know* is highly speaker-oriented: its purpose is to create salient coherence relations and help the speaker when s/he is having difficulties in planning the utterance.

Thus, Müller (2005) and House's (2009) point of views are slightly different, House (2009) claiming that *you know* functions primarily on textual level and Müller (2005) arguing that it functions on both levels. Despite the differences between these two studies, both are an essential point of reference for this thesis when analysing *you know* in *English as Lingua Franca (ELF)* interaction. Furthermore, the differences show that there is a need for this type of study to define the functions *you know* may have in ELF interaction.

In this thesis, ELF speakers are considered as one collective group that a) speak English as a foreign language, b) do not share the same mother tongue and c) use English as a language of communication (Firth 1996: 240). The data of this thesis is from the corpus of English as Lingua Franca in Academic Settings, i.e., the ELFA corpus (ELFA 2008). It is a corpus of face-to-face academic interactions spoken by ELF speakers, and the data in the ELFA corpus is not elicited but occurs naturally. The size of the ELFA corpus is one million words with 650 speakers and 51 different first languages. (ELFA 2008.)

The data used in Müller (2005) and House's (2009) studies are similar to the ELFA corpus in that the speakers have an academic background, but the size of their self-collected corpora is smaller than the size of the ELFA corpus. Furthermore, Müller (2005: 31–32) uses the Giessen-Long Beach Chaplin Corpus (GLBCC) based on silent movie narratives, and House (2009: 178) uses a corpus based on informal conversations about short provocative articles. In other words, the data is quasi-authentic in both corpora, which means that the conversations are not completely natural but elicited by the researchers. In addition, although the interactants in House's (2009) study have many first languages, the speakers in Müller's (2005) data are only Germans and Americans, which is why Müller's (2005) data cannot be referred to as ELF interaction in its purest sense (cf. Firth 1996: 240). Therefore, there is a need to study *you know* by using a large spoken ELF corpus that, in turn, may bring new information to the pre-established functions of *you know*.

As the data for this study is from a corpus, a corpus-driven approach is utilised. According to Tognini-Bonelli (2001: 84), a corpus-driven approach takes into account all examples of the analysed linguistic item without any modifications or limitations. Therefore, all instances of *you know* spoken by ELF speakers in the ELFA corpus are taken into account in the analysis.

To conclude, using Müller (2005) and House's (2009) studies as a baseline, the primary aim of this thesis is to study how ELF speakers use the discourse marker *you know* in face-to-face academic interaction. Moreover, the aim of this thesis is to see how the results of *you know* in the ELFA corpus correlate with the aforementioned studies on non-native English interaction. Thus, the research questions are as follows:

- 1) How is *you know* used as a discourse marker by ELF speakers?
- 2) How do the functions of *you know* in the ELFA corpus correlate with previous studies, particularly Müller (2005) and House's (2009), on *you know* in non-native English interaction?

The first research question is answered by analysing the different functions *you know* as a discourse marker has in the ELFA corpus, whereas the second research question is answered by comparing the results of the analysis with the results of Müller (2005) and House's (2009) studies. In addition, if Müller (2005) and House's (2009) functions of *you know* do not correspond to the functions found from the ELFA corpus, previous studies on native English speakers' usage of *you know* are also used as a reference point.

In Section 2, the key aspects and terminology regarding the discourse marker *you know* are introduced and defined, in addition to the pre-established functions of *you know* in non-native English interaction. In Section 3, the data used for this study as well as the method of analysis are introduced, whereas Section 4 consists of the analysis. In Section 4.1, the functions that *you know* has as a discourse marker in the ELFA corpus are analysed and described, and in Section 4.2, the results of

the analysis are compared to previous studies. Lastly, conclusions of the results are described in Section 5.

2. Theoretical background

Before *you know* as a discourse marker in academic ELF interaction can be defined, not to mention analysed, some relevant terminology need to be explained. For example, the term *discourse marker* itself has had many slightly different definitions over the years. Therefore, relevant terminology is defined in Section 2.1. In section 2.2, on the other hand, the pre-established functions of *you know* in previous literature are introduced. These pre-established functions act as a framework in analysing the instances of *you know* as a discourse marker that occur in the ELFA corpus.

2.1 Relevant terminology

As the term itself suggests, discourse markers occur in *discourse*. Schiffrin, Tannen and Hamilton (2015: 1) define discourse as “(1) anything beyond the sentence, (2) language use, and (3) a broader range of social practice that includes non-linguistic and non-specific instances of language”. Therefore, *academic discourse* can be defined as discourse occurring in an academic setting that consists of the use of language, the situation it is spoken in, as well as the social aspects that are explicitly or implicitly present in academic speech events.

Discourse marker is a term not simply defined for many reasons. Although most scholars agree on discourse markers having a pragmatic function in an utterance, they seldom agree on anything else (Müller 2005: 1). For example, scholars use different terms in defining the same lexical expression (Schourup 1999: 227), which is why *you know* has been called a *discourse marker* (e.g. Schiffrin 1987), *pragmatic particle* (e.g. Östman 1981), *pragmatic expression* (e.g. Erman 1987), *pragmatic*

marker (e.g. Erman 2001), etc. In addition, the terms are not necessarily interchangeable as their definitions vary between different scholars (Jucker & Ziv 1998: 2; see also, e.g., Fraser 1996, 1999; Schiffrin 1987). Moreover, different scholars give different requirements for these lexical expressions, which means that discourse markers have been given numerous functions in previous literature (see, e.g., Fraser 1996, 1999; Schiffrin 1987). That is why discourse markers have also been called to function, e.g., as *discourse connectors*, *confirmation-seekers*, *hesitation markers*, *fillers* or *repair markers* (Jucker & Ziv 1998: 1).

Therefore, defining how the term *discourse marker* is utilised is important. This thesis follows Müller's (2005: 27, original emphasis) definition of discourse markers as linguistic elements:

- which are, as a group, difficult to place within a traditional word class,
- which are *syntactically* optional,
- which may occur at the beginning, middle, or end of a discourse unit or form a discourse unit of their own,
- which have little or no semantic meaning in themselves,
- which are multifunctional,
- which occur in oral rather than written discourse. If they are found in both, they often if not always assume functions in the oral medium that go beyond those they have in the written medium

Hence, if *you know* in the ELFA corpus meets the abovementioned requirements, it is treated as a discourse marker. The emphasis is particularly on *you know* having little or no semantic meaning and its syntactic optionality: if *you know* can be omitted without harming the overall meaning of the message and the syntactic structure of the sentence, it is considered as a discourse marker.

Since the data for this thesis comes from the ELFA corpus, the term *English as Lingua Franca* (ELF) ought to be defined as well. In this thesis, the definition of ELF follows Firth (1996: 240) who views ELF as a “‘contact language’ between persons who share neither a common native tongue nor a common (national) culture, and for whom English is the chosen foreign language of communication.” In addition, the *A* in ELFA refers to academic settings, which suggests that although

non-native speakers' level of English is not similar to native speakers' level, they can nevertheless be regarded as fluent speakers of English because they use it as a medium in academic discourse.

Similar to the definition of discourse markers, the definition of ELF also sets some conditions for the speakers whose usage of *you know* is taken into account in the analysis. In other words, native speakers as well as speakers who have stated English as their second language are excluded from the data, as they do not meet the conditions of an ELF speaker. Moreover, speakers whose mother tongue is unknown are also excluded from the data as one cannot be certain that they fit in with the definition of an ELF speaker.

2.2 Defining the functions of *you know*

Many scholars have studied how native English speakers use *you know* as a discourse marker (e.g. Östman 1981; Holmes 1996; Schiffrin 1987; Erman 1987, 1992, 2001; Redeker 1990; Macaulay 2000), and according to Müller (2005: 147), it has gained almost 30 different functions over the years. Therefore, Müller (2005: 147) calls *you know* “one of the most versatile and notoriously difficult to describe”. Moreover, it seems that recently the interest has expanded on studying non-native English speakers – especially learners – and their use of *you know* as well. For example, Fung and Carter (2007) compare the use of *you know* between Hong Kong learners and native English speakers, and Polat (2011) studies the use of *you know* by immigrant second language learners of English.

Müller (2005) and House's (2009) data, on the other hand, consist of non-native university students who are relatively *fluent* speakers of English. Moreover, their level of English as well as their academic background are similar to the ELF speakers in the ELFA corpus, which also means that they are suitable studies for reference and comparison. Consequently, as there are approximately 30 different functions accorded to *you know*, the functions introduced in the following subsections are mostly based on the pre-established functions by Müller (2005) and House (2009).

The functions of *you know* are divided into ten different functional categories in this thesis, and although they are mostly based on Müller (2005) and House's (2009) pre-established functions, their names primarily stem from Müller's (2005: 157) categories. Table 1 shows Müller's (2005: 157) original categories on the left and the adapted categories used in this thesis on the right:

Table 1. Comparison of the categories of *you know*

| Müller | Adapted version |
|---------------------------------------|---------------------------------------|
| Textual level | |
| marking lexical or content search | marking lexical or content search |
| marking false start and repair | marking repair |
| marking approximation | - |
| introducing an explanation | introducing an explanation |
| quotative <i>you know</i> | quotative <i>you know</i> |
| - | linking propositions together |
| - | highlighting certain elements |
| Interactional level | |
| "imagine the scene" | securing comprehension |
| "see the implication" | |
| appeal for understanding | appeal for understanding |
| reference to shared knowledge | reference to shared knowledge |
| acknowledge that the speaker is right | acknowledge that the speaker is right |

As table 1 shows, most names are relatively similar, but there are some exceptions, additions and omissions as well. For example, Müller (2005: 162–164) discusses *you know* functioning as a marker of approximation, but as it does not occur in the ELFA corpus, it will not be included in this thesis.

Furthermore, some of the functions listed above might occur in Müller's (2005) study but not in House's (2009), and vice versa. Moreover, it became clear during the analysis that Müller (2005) and House's (2009) functions are insufficient in defining the functions of *you know* in the ELFA corpus (see, e.g., 2.2.5 *Linking propositions together*). To compensate this, other previous studies on native English speakers' use of *you know*, Erman's (2001) study in particular, were also used as a reference point.

Table 1 also shows the discourse marker *you know* to function on two levels: textual and interactional (see also Müller 2005: 30–31). Müller (2005: 30–31) observes that on textual level, *you know* focuses “on the lexical expressions and propositional content expressed in units of various length, from single words or phrases to a sequence of utterances”, whereas on interactional level, it focuses “on the relationship between speaker and hearer.” In other words, on textual level *you know* is directed towards the textual aspects of the utterance, whereas on interactional level, it functions on a more abstract level trying to gain the addressee's involvement. House (2009: 171), on the other hand, disagrees with Müller (2005) and many other studies that consider *you know* functioning on interactional level, arguing that ELF speakers do not use *you know* to create interactional relations but use it primarily as a subjective device, i.e. that it functions only on textual level.

Before moving on to defining the functions of *you know*, there are some relevant terms used in analysing *you know* in previous literature as well as in this thesis that need to be shortly addressed. For example, the position of *you know* in an utterance or a *turn-unit* (later *turn*) often determines what kind of pragmatic function it has. According to Sacks, Schegloff and Jefferson (1974: 702), *turns* are a constructional component of speech that can be divided into different unit-types, including “sentential, clausal, phrasal, and lexical constructions”. Therefore, a turn can consist of a sentence, *I'm here because I didn't know where else to go*, a clause, *because I didn't know*, a phrase, *didn't know*, or a lexeme, *know*. The notion of clause, however, is not as strict in spoken language as it is in

written language. For example, conjunctions such as *and*, *but*, *because* or *so* do not automatically begin a subordinate clause in spoken language, but they can begin a main clause as well.

Turn-taking is another important aspect that aids in defining the functions of *you know*. Turn-taking deals with speaker-selection during conversation. Sacks et al. (1974: 705) explain that the next speaker is selected either by the current speaker or the next speaker him/herself in a conversation. If the current speaker selects the next speaker, indications of turn-transition are, e.g., silence or a question that naturally leads to the next person taking the turn, i.e. giving an answer (Sacks et al. 1974: 705). Other similar sequencing *adjacency pairs* to question–answer are, e.g., greeting–greeting and offer–acceptance/refusal (Schegloff and Sacks 1973, quoted in Schegloff 1988: 109).

House (2009) and Erman (1987, 2001) also use the terms *theme* and *rheme* in defining the position of *you know* in an utterance. Halliday and Matthiessen (2004: 169) define *theme* on textual level as the part of turn that consists of the message, i.e. the topic of the clause. *Rheme*, on the other hand, is what elaborates the theme. According to Halliday and Matthiessen (2004: 64–65, 169), the theme usually comes first in a clause, often being a noun phrase that occurs before the main verb, like in the example they provide: “[Theme:] *It* [Rheme:] *means more often than not.*” The terms *theme* and *rheme* are especially present in the function of *you know* highlighting certain elements (see 2.2.6 *Highlighting certain elements*).

Lastly it must be stated that Müller (2005) and Erman (2001) occasionally use the speaker’s intonation contour to define in greater detail the functions of *you know*. For example, Müller (2005: 184) uses intonation contour to define if the speaker states, asks or demands the addressee to agree with him/her (see 2.2.10 *Acknowledge that the speaker is right*). However, as this thesis utilises a corpus-driven approach, its aim is to rely on the ELFA transcripts, not the recordings, which means that the recordings of the ELFA corpus were only used if they were absolutely necessary. This was the case, e.g., with some of the instances of *you know* when they were divided into discourse markers and non-discourse markers.

In the next subsections, the ten adapted categories presented in table 1 are further explained. They are based on previous literature, mostly Müller (2005) and House (2009), as well as on the analysis of the ELFA corpus. The first six categories function on textual level, whereas the remaining four categories function on interactional level.

2.2.1 Marking lexical or content search

The first function of *you know* as a discourse marker on textual level is called *marking lexical or content search*, a name given by Müller (2005: 158). The primary indicators of this function are pauses as well as verbal hesitations, such as *er* or *um* (Müller 2005: 158–159). Some previous studies suggest that the place of a pause may determine whether the speaker is searching for a lexeme or propositional content (see e.g. Östman 1981; Erman 1987), but Müller (2005: 160) argues that in many a time the distinction is impossible to make. This is the case in example (1) from Müller (2005: 159):

- (1) um ... (1.1) the .. the captain of the ship accuses him of being a pickpocket, (H) .. and um .. the scene -- the the situation is then revealed to the young girl who realizes that Charlie's put the money in her pocket, and (H) um (H) whether -- uh I don't think she .. *you know*, .. really considers whether he taken it in in the first place,

Östman (1981: 30–31) suggests that pauses are not necessary to mark textual hesitation but when they do, their place in the utterance is meaningful. Östman (1981: 29–30) claims that if *you know* is followed by a pause, the speaker is searching for a word, and if a pause (potentially) precedes *you know*, the speaker is searching for propositional content. However, in example (1) – and in many other cases in Müller's (2005: 160) data – *you know* is preceded and followed by a pause, which makes Östman's definition difficult to confirm. Thus, Müller (2005: 160) suggests that 'marking lexical or content search' is a category, which is "a continuum of instances with clear cases of lexical

search at one end ..., apparent content search at the other, and cases with elements of both in between.”

House (2009: 171) also discusses lexical and content search but uses the phrase “fumbling for words” to describe how the primary function of *you know* in this sense in ELF interaction is to stall for time when the speaker is hesitating what to say next. In other words, House (2009: 186) does not differentiate lexical search from content search but states that *you know* is used when the speaker “hesitates, cannot find the right words, [or] fumbles for the appropriate word or formulation” to signal “current planning difficulties”. According to House (2009: 186), the position of *you know* in these cases is in mid-utterance, usually occurring inside nominal, verbal and adverbial groups – a term coined by Halliday (1985) for systemic-functional grammar. House (2009: 171, 187) also emphasises that in ELF interaction, this function of *you know* is not only to show hesitation but also to create coherence by acting as a focusing device when the speaker is, for one reason or another, not able to continue.

Creating coherence is also noted by Erman (2001: 1340) who in her study on native English speakers notices how native speakers alike use *you know* when searching for a correct structure or stalling for time to properly continue their utterance. Erman (2001:1346) also offers a more detailed position for *you know* when it functions as a marker of lexical or content search. According to Erman (2001: 1346), when *you know* occurs “within the phrase after a determiner”, the speaker is usually searching for words, and when *you know* occurs “after a con/disjunct at the beginning of the clause”, the speaker is usually searching for content. Unfortunately, Erman (2001: 1344–1345) does not define the terms any further, but gives two examples where in the first example *you know* is surrounded by the preposition *on* (a determiner) and in the second example *you know* is preceded by *cos*, which is a short form of the conjunction *because* (a conjunct).

Determiners are words that make a noun phrase either definite (*the, my, those, which*) or indefinite (*a, any, other, what*) (McArthur 2003b, s.v. *determiner*). A conjunct, on the other hand, is

one of the many terms used for words that function as a linking device (e.g. see Liu 2008: 492). Liu (2008: 492) points out that often the term *conjunct* is used specifically to adverbials functioning as a linking device. These are, e.g., *furthermore*, *also*, *next*. However, based on Erman's (2001: 1345) example, it can be assumed that she utilises the term *conjunct* to include all linking devices, conjunctions included (cf. Liu 2008: 492). *Disjuncts*, consequently, are sentence adverbials that either express the speaker's attitude (e.g. *frankly*) or evaluate what is being said (e.g. *perhaps*) (McArthur 2003a, s.v. *adverbial* 2).

2.2.2 Marking repair

The second function of *you know* on textual level is called *marking repair*, a name adapted from Müller's (2005: 160) category "marking false start and repair". According to Müller (2005: 160–161), repairs must occur within one syntactic structure, and they consist of a) cases where a complete or unfinished word is substituted with another one after *you know* or b) cases where nothing is repaired but only repeated. Müller (2005: 161) also differentiates these two types of repairs into repairs and false starts, respectively. However, such a distinct differentiation is not used in this thesis, but instead it follows Fox and Jasperson's (1995: 80) view that false starts are a subtype of repair. This is also why the category is simply called 'marking repair'.

However, omitting "false start" from the name does not mean that false starts are not included in this category. Example (2) from Müller (2005: 161) demonstrates such a repair (i.e. a false start):

- (2) he ... he was sort of dressed like -- *you know*, like you knew he didn't have no .. not too much money

In example (2), the speaker cuts off in the middle of the utterance (indicated by two dashes), adds *you know* and continues after repeating the word *like* (Müller 2005: 161). Thus, nothing is repaired but

merely repeated. It must be noted, however, that repetition of one or more words is not only a feature of false starts but it may occur in other types of repairs as well (Müller (2005: 161). Moreover, it is the context, not repetition, that defines the function of *you know*, as repetition of words may also occur when the speaker is introducing an explanation (see 2.2.3 *Introducing an explanation*).

Müller (2005: 161) takes a strict approach to repairs stating that she does not take into account cases where the syntactic structure of the utterance changes after *you know*. This is because Müller (2005: 162) only had one such instance in her data. Holmes (1986: 11) and Erman (2001: 1345), on the other hand, take into account cases where the speaker uses a new syntactic structure after *you know*. Erman (2001: 1345) refers to these as restarts. The excerpt below from the ELFA corpus serves as an example of a restart:

- (3) <S1> and i don't want *you know* i have s- problem of of seeing some no locating things...
(CDIS01A)

S1 begins the sentence with a negative sentence structure *I don't want*, but stops in mid-utterance and after *you know* makes a repair that changes the syntactic structure as well as most of the words: *I have s- problem of of seeing*. Because of cases like example (3) that occur in the ELFA corpus, a less strict approach is utilised in this study.

House's (2009) approach is rather different to Müller's (2005) although they both agree that *you know* as a marker of repair functions on textual level. House (2009: 186) states that in addition to lexical and content search, fumbling for words includes the speaker's attempt to repair his/her misstep with the help of *you know*. In other words, she does not make a distinction between lexical or content search and repair, but classifies both of them as fumbling for words. Thus, the position of *you know* is the same as in the category 'marking lexical or content search': in mid-utterance as well as inside nominal, verbal and adverbial groups (House 2009: 186). Similar to Müller (2005), House (2009: 186) also notes that *you know* functions on textual level – or in her words, “at a more local,

micro-level.” Although a unified account is undeniably desirable (cf. House 2009: 188–189), based on the analysis of the ELFA corpus, a division into lexical or content search and repair seems also justifiable.

2.2.3 Introducing an explanation

According to Müller (2005: 167), *introducing an explanation* functions on textual level, and the purpose of *you know* is to point out “ideational relationships between statements or concepts which precede it and those which follow it.” According to Halliday and Matthiessen (2004: 309), there are three different functional components to construct meaning, *ideational metafunction* being one of them. Ideational metafunction means that the meaning of a clause is seen as a representation of how the speakers experience the world on the one hand, and what kind of logical generalisations they can make about their world on the other hand (Halliday and Matthiessen 2004: 309–310). Based on the examples Müller (2005: 165–167) provides, she most likely uses “ideational relationships” to refer to the speakers’ experiences of the world, or in her interactants’ case their experiences of the silent movie, and how *you know* connects and explains the clauses that represent the speakers’ experiences.

House (2009: 182–185), on the other hand, does not explicitly refer to *you know* as a marker introducing an explanation, but she also mentions *you know* marking ideational relations with and without conjunctions, such as *and*, *but* and *because*. In addition, House (2009: 184) notes that as a relational phrase, *you know* “is used to indicate implicit conjunctive or coherence relations of addition (elaboration, extension, expansion in the Hallidayan sense), of contrast, opposition and concession as well as causal relations WITHOUT the overt presence of such cue words as *and*, *but* and *because*.” In other words, *you know* functions as a substitute for conjunctions and other words that create coherence, which in turn is similar to Müller’s (2005) definition of *you know* introducing an explanation.

The key point of this function is that the speaker voices an opinion but feels the need to express it differently in order to avoid misinterpretation or misunderstanding (Müller 2005: 166). In other words, the speaker is helping the addressee to interpret the message, as in example (4) from Müller (2005: 167):

- (4) [[but then again there's nothing else]]. *you know* it's like -- there was no other kind of entertainment though.

In example (4), the speaker voices an opinion *there's nothing else*, but decides that an explanation is needed to make the meaning of his argument plainer (Müller 2005: 166).

Müller (2005: 167) notes that there are different types of explanation that should be taken into account. These are, e.g., clarification, exemplification and giving additional background information (Müller 2005: 165–167, see also Holmes 1986: 11; Crystal 1988: 47; Erman 1987: 114). In addition, cases in which the speaker introduces, in Erman's (2001: 1342) words, "a change of information content, frequently correcting or modifying previous discourse" are also categorised as 'introducing an explanation'. Example (5) from Erman (2001: 1342) demonstrates such a case:

- (5) /.../ and he's sort of next one, *you know* next senior one after Hart.

In example (5), the speaker adds the word *senior* to specify the previous discourse (Erman 2001: 1343). Moreover, specification often occurs with repetition (Erman 2001: 1343), which is the case in example (5) as well.

Erman (2001: 1342) explains that the difference between a repair and introducing a change of information content is that in repair, the speaker makes a restart before finishing the syntactic structure of the previous utterance. Thus, as *you know* occurs after the syntactically completed

structure *sort of next one* in example (5), it is, according to Erman (2001: 1343), considered a modification that “introduces specification of previous discourse”.

Parenthetic comments are also a way to offer additional information, and Erman (2001: 1344) observes that when the speaker assumes that additional information is necessary for the addressee to understand the narrative, *you know* functions as a marker to introduce parenthetic comments. In addition, unlike with instances where the speaker introduces a change of information content, parenthetic comments may occur in mid-structure (Erman 1992: 223). Erman (2001: 1344) gives an example:

- (6) <1> /.../ Shelley, come round to me right, and she was, stroking Dempsey and he walked past wagging his, *you know* when they put the tail down [and]
 <2> Yeah.
 <1> (continues)

In example (6), instead of finishing the previous utterance and the transitive verb *wag* with an object, the speaker offers additional information *when they put the tail down* that s/he thinks is relevant for the addressee to know.

Offering background information is a function that seems to be difficult for Müller (2005: 166, 186) to categorise as she mentions it in two different categories, and in both cases uses Erman (1987, 1992, 2001) and Östman (1981) as a reference. At first, *you know* offering background information appears under the category ‘introducing an explanation’, and Müller (2005: 166) offers background information as an example for the superordinate *amplification*, a term she uses to describe a form of explanation. Müller (2005: 166) states that the instances in which *you know* functions as an amplification are rare in her data but gives an example of such an instance nevertheless – even referring to it as “a parenthetic comment”. The second time, however, it appears under the category ‘various functions’ (Müller 2005: 186). ‘Various functions’ consists of instances whose functions are clear but the number of occurrences is too scarce in Müller’s (2005: 186) data. This time, Müller

(2005: 186) claims that there is only one such instance where *you know* functions as an introductory to “relevant background information, for example in a parenthetic comment”, and gives that instance as an example.

In this thesis, offering background information is considered a form of explanation because its purpose is to help the addressee to interpret the message. Like in example (6), the speaker is not sure that the addressee has all the information needed to understand the narrative, offers additional background information in a parenthetic comment and then continues with the argument or narrative. That is why offering background information is categorised as ‘introducing an explanation’ in this thesis.

2.2.4 Quotative *you know*

The function of *you know* as a quotative device is a function established by Müller (2005) but not by House (2009). However, as Müller (2005: 167–168) notes, it is not a common function in her data. It is not a common function in the ELFA corpus, either, but worth a category nevertheless.

Reported speech can be defined into direct-speech reports and indirect-speech reports. Erman (2001: 1344), for example, compares the function of *you know* between the speaker’s own speech and (directly) reported speech to quotation marks. Therefore, if direct-speech report is comparable to speech inside quotation marks, it can be argued that indirect-speech report is comparable to paraphrased speech. However, as Müller (2005: 170) observes, direct-speech report may also be hypothetical speech as speakers in her data reported a character’s lines as direct speech although the movie is silent. Thus, direct-speech reports do not necessarily have to be word-for-word in order to be viewed as direct.

Müller (2005: 168) only takes into account cases of direct-speech reports that follow *you know*, and pays attention to the voice quality of how the quotation is said. He and Lindsey (1998: 143), on

the other hand, include both directly and indirectly reported speech in their data stating that the report *you know* introduces offers “new, important, salient information” to discourse. In addition, Erman (1992: 221) takes into account cases where *you know* comes after reported speech. When *you know* occurs before reported speech, it may appear with a quotative verb, such as *think*, *say* or *ask*, as well as with the construction *BE + like* (the capitalised *BE* representing all the forms of the verb), which functions the same way as quotative verbs (Müller 2005: 169).

Example (7) is from the ELFA corpus and it demonstrates *you know* being used as an enquoting device before direct-speech report. *You know* is also preceded by a quotative verb *said*:

- (7) <S1> he said hey this is very dangerous place i said why we're in united states what's wrong
he said *you know* this place oh <FOREIGN> pericoloso </FOREIGN> (ULECD030)

As the excerpt shows, using *you know* to introduce reported speech is not necessary but using it does, however, create coherence and highlight the reported speech occurring after *you know*. Thus, as *you know* is directed towards the textual aspects of the utterance, quotative *you know* operates on textual level (see also Müller 2005: 170).

Although Müller (2005: 168) focuses on the speaker’s voice quality when defining quotative *you know*, it is not done in this thesis. As mentioned before, the aim of this thesis is to be corpus-driven, i.e. to rely on the transcript of the corpus, not the recordings. The received recordings were only used if they were absolutely necessary, and in the case of *you know* as an enquoting device, the transcript was sufficient.

Lastly, it must be stated that ‘quotative *you know*’ is either not a typical function among ELF speakers or it is not seen as a function that is common enough to be given a category of its own. For example, none of Müller’s (2005: 170) instances of quotative *you know* come from non-native speakers, and House (2009), whose data consists solely on ELF speakers, does not even mention it

in her study. However, as there are clear cases in the ELFA corpus where *you know* introduces reported speech, it cannot be omitted from this thesis.

2.2.5 Linking propositions together

Linking propositions together is a function that is not mentioned by Müller (2005). House (2009), on the other hand, observes that *you know* is used as a relational phrase “to make implicit coherence relations more explicit”, which could be considered as linking propositions together. However, the implicit coherence relations House (2009: 184) refers to are primarily cases where *you know* acts as a substitute for conjunctions *and*, *but* and *because*, which means that *you know* is performing the additive, adversative and causal functions instead of the aforementioned conjunctions, respectively (cf. 2.2.6 *Highlighting certain elements*).

Erman (1987: 82–83) similarly observes that *you know* sometimes functions as “a connector” indicating causal relations between two propositions, and based on the examples she provides, the conjunction *because* is not present in these cases. Later on, Erman (2001: 1342) compares *you know* to conjunctions in general, and proposes that one of the functions of *you know* is to set a link between propositions, such as moves between arguments. However, in this case, most of the examples Erman (2001: 1343–1344) provides include conjunctions. Therefore, it could be argued that Erman does not consider the function of *you know* to be merely a booster of or a substitute for a conjunction, but that when the two occur together, they also function together (see, e.g., Erman 2001: 1343; *so y’know*).

Example (8) is the only example from Erman (2001) where *you know* as a connector occurs without a conjunction. Example (8) demonstrates a move between arguments (Erman 2001: 1343):

- (8) /.../ they did it in a completely, slapstick farce way, *you know*, the the men who were dressed up supposed to be women had great big balloons and, had rosy red cheeks and wigs and things /.../

Erman (2001: 1343) explains that in argumentative discourse, *you know* often occurs between “the speaker’s position and the backing up of it”. That is also the case in example (8): the speaker voices an opinion *they did it in a completely slapstick farce way* and elaborates it in the utterance following *you know*. Moreover, it could be also argued that in example (8) *you know* is substituting the conjunction *because* (see House 2009: 184) or a colon (:) that can have analogous functions to *because* in written text. This, moreover, suggests that on Müller’s (2005) terminology, *you know* linking propositions together operates on textual level.

‘Linking propositions together’ is very similar to the category ‘introducing an explanation’ as both of them occur between statements. However, Erman (2001: 1343) notes that when *you know* links propositions together, it is not used to aid the addressee in interpreting the message but to introduce a proposition or mark a connection between propositions that are “largely independent”. Thus, in example (8), the cause of elaboration is not because the speaker is afraid of being misinterpreted as it would be in introducing an explanation. Rather, the speaker has stated his/her opinion in the first utterance and backs it up in the second that follows *you know*.

2.2.6 Highlighting certain elements

Highlighting certain elements is a function established in House (2009) but not categorised by Müller (2005). House (2009: 181) notes that *you know* occurs in close vicinity of conjunctions *and*, *but* and *because*, and infers that in these cases the function of *you know* is to highlight and make more explicit the relations these conjunctions express – addition, opposition and causality, respectively. Thus, the function of *you know* is to boost or reinforce, and therefore highlight, the relations these conjunctions create (House 2009: 184). This, as House (2009: 183) suggests, is a speaker-oriented function, which in Müller’s (2005) terminology operates on textual level.

Erman (2001: 1342), on the other hand, argues that the primary function of *you know* on clause level is to aid the addressee to interpret the received message. Thus, in Müller's terms, it functions more on interactional level. One of those forms of aid is *you know* being used to highlight certain elements (Erman 2001: 1342). 'Highlighting certain elements' is a function that can also be defined with the help of Halliday's (1985; also Halliday and Matthiessen 2004) systemic-functional grammar. For example, Erman (1987: 130–131) defines *you know* as an introductory device and explains that the function of *you know* is to organise the thematic structure of the sentence. Halliday and Matthiessen (2004: 64) explain that the thematic structure of an utterance consists of the theme and the rheme, where the theme is "the point of departure of the message" and the rheme the part following the theme that also develops the theme further.

According to Erman (1987: 130, 132), the position of *you know* highlighting certain elements is sometimes after a subject noun phrase or fronted adverbial (i.e. an adverbial that precedes the subject). Erman (1987: 131–132) observes that when *you know* occurs after the subject noun phrase, it serves to "round off the theme", whereas after an adverbial phrase, it creates a link back to the previous discourse. Example (9) is of *you know* occurring after the adverbial phrase *the next minute* (Erman 2001: 1342):

- (9) /.../ and we, we all buy, we buy chips yeah, and the next minute, *you know*, we're all walking into the arcade all these girls just come up to us and start taking chips /.../

Erman (2001: 1343) determines that in example (9) the function of *you know* is to mark and possibly highlight the fronted adverbial *the next minute* that acts "as a scene setter" for the following course of events. At first, it might seem that Erman (2001) is suggesting that the fronted adverbial *the next minute* is the most important piece of information the addressee needs to interpret the message. However, in an earlier study, Erman (1987: 130) states that when a fronted adverbial precedes *you know*, its function is to introduce the upcoming argument – or in this case the ensuing course of events. Therefore, it can be argued that *you know* is not merely highlighting what comes before it, but more importantly, what comes

after it, i.e. the rheme or new information in Halliday's (1985) terminology. House (2009), on the other hand, would likely emphasise the presence of the conjunction *and* in example (9).

House (2009: 183) also notes that some functions of *you know* can be defined in Hallidayan sense to theme and rheme or given and new. Similar to Erman (1987), House (2009: 183) states that *you know* seems to create a link between the theme and what comes after it, and paraphrases the theme as "this is the heading to what I am saying". However, House (2009: 183, 187–188) criticises that *you know* does not introduce anything given, i.e. information that the addressee is able to recover, even though the core meaning of *you know* insinuates otherwise. In her later study, Erman (1992: 222; 2001: 1342), on the other hand, notes that with *you know* the speaker urges the addressee to accept part of the information that typically occurs before *you know* as given or common ground. Thus, *you know* is not used to address given or retrievable information, but it is used to persuade the listener that such information exists – although it necessary does not. This function, consequently, leads towards another function, 'reference to shared knowledge' (see 2.2.9 *Reference to shared knowledge*).

2.2.7 Securing comprehension

All the functions that are defined henceforth perform on interactional level. When *you know* functions on interactional level, it is used to elicit some kind of involvement or co-operation from the addressee (Müller 2005: 171). Eliciting involvement, co-operation or any other mode of intersubjectivity is, however, a function not supported by House (2009). In fact, House (2009: 190) strongly argues against it, claiming that ELF talk is primarily self-centered and *you know* is used to aid the speaker, not the addressee, to create textual connections or help during planning difficulties. However, as the analysis of *you know* in the ELFA corpus later on shows, *you know* can also have functions on interactional level.

The first function of *you know* on interactional level is called *securing comprehension*. It is a term adapted from Erman (2001) and it combines two of Müller's (2005: 171–177) functions that she paraphrases as “imagine the scene” and “see the implication”. The difference between the two is slight and the reason why Müller (2005: 171) has divided them into two different categories is due to the type of corpus she uses, i.e. a corpus of movie narratives, in which case it might be important to make a clear distinction between the two. However, the baseline of “imagine the scene” and “see the implication” is the same, which is why they are combined in this study.

Müller (2005: 175) explains that the basis of both “imagine the scene” and “see the implication” is to offer a request for the addressee to get involved in the narrative either mentally or actively. If the desired involvement is an active, oral response, *you know* may also have a confirmation-seeking function that naturally leads to turn-yielding (see also Östman 1981: 27; Erman 2001: 1345). Moreover, Crystal (1988: 47) aptly compares *you know* seeking confirmation to a tag question when it occurs at the end of a sentence.

“Imagine the scene” is a function in which the speaker is describing a scene in narrative discourse, and with *you know* the speaker is either asking the addressee to visualise the scene in his/her head or being sure that the addressee is able to do so (Müller 2005: 171). “See the implication”, on the other hand, is a function in which the speaker implies something and uses *you know* to ask, check or make sure that the addressee is able to understand the implication (Müller 2005: 175). Both functions are similar to Erman's (2001: 1346) comprehension-securing function. Erman (2001: 1346) explains that *you know* is commonly used in the social domain – in Müller's (2005) terminology on interactional level – to make sure the addressee “has correctly understood specific references made in the text, usually about people but also objects and other phenomena”. In other words, *you know* asks the addressee to connect the dots between what is implied and what is actually said, as well as to check that s/he is able to do so (see also Jucker & Smith 1998: 196).

Example (10) from the ELFA corpus shows how the difference between “imagine the scene” and “see the implication” is not always easily distinguished:

- (10) <S2> cadmium is something that you need to be particularly worried about it's in your mobile phone and now cadmium gives you kidney problems that's why i see many people standing like this *you know* <SS> @@ </SS> (ULEC01B)

S2 is explaining the possible side effects of cadmium to kidneys and demonstrates how people are nowadays standing due to these problems. As the demonstration itself is visual, S2 is probably sure that the addressees are able to picture the scene, yet *you know* is still added at the end of the utterance to get the listeners involved (which they do by laughing). Moreover, although the connection between kidney problems and standing in a particular position seems apparent, the implied connection is emphasised with *you know*. However, it can be argued that the important part in example (10), or in any other case alike, is not whether it is about imagining the scene or seeing the implication, but that in the end the function of *you know* is to secure the addressees' comprehension.

Securing comprehension is closely connected to speech acts. In *A Dictionary of Media and Communication* (Chandler & Munday 2016, s.v. *speech act*, original emphasis), *speech act* is defined as

Any goal-directed action performed with words in interpersonal communication, defined primarily with reference to the speaker's intentions and the effects on the listener(s). The term was introduced by John Austin and is also associated with John Searle in an analytical approach called **speech act theory**.

Thus, when *you know* is used to secure the addressee's comprehension by, e.g., asking the addressee to understand the speaker's implication, the purpose of *you know* is to underline the speaker's intention behind the utterance in addition to its apparent meaning.

Müller (2005: 171) uses the level of intonation to identify whether the speaker is asking, checking or being sure that the addressee is able to “imagine the scene” or “see the implication”.

However, as intonation is not marked in the transcript of the ELFA corpus, it is not taken into account in this study. Moreover, in the majority of cases, the transcript was sufficient in defining *you know* as securing comprehension, and in the minority of cases where it was not, the recordings were used to ascertain the function of *you know*.

2.2.8 Appeal for understanding

According to Müller (2005: 181), *appeal for understanding* is a category similar to ‘lexical or content search’, because in both cases the speaker is unsure of what s/he is saying or is about to say next. What differentiates the two, however, is the level on which they function. In lexical or content search, *you know* functions on textual level as the speaker is searching for a proper word or content. However, in appeal for understanding, the speaker cannot or does not try to find the words but appeals to the listener to fill in the missing pieces. Thus, the speaker is inviting the addressee’s involvement, which makes *you know* to function on interactional level. (Müller 2005: 181.) Example (11) is from Müller (2005: 181):

- (11) A: .. remember [this] dress and this &
 B: [yeah],
 A: & um ... (1.1.) <L2 Kopftuch L2>?
 [[<L2 Kopftuch L2>]]?
 B: [[yeah I don’t know]].
 [yeah].
 A: [this] *you know*,
 and he he was falling down,

In example (11), the speaker is unable to come up with an English word for *scarf* and the function of *you know* is to appeal for understanding and the addressee to fill in the gap (without success) (Müller 2005: 181).

In example (11), the appeal for understanding involves lexical choices (Müller 2005: 182). However, it can also occur in argumentative discourse when the speaker is either insufficient in expressing an opinion or when his/her knowledge about the subject is inadequate (Müller 2005: 182–183). Example (12) demonstrates the latter. Before the excerpt in example (12), speakers A and B discuss a character's clothes with speaker A arguing that they do not fit the situation presented in the movie (Müller 2005: 182–183). Speaker B responds (Müller 2005: 183):

- (12) B: mhm. ... it's true but -- I don't know .. I don't know *you know*, I never look for the clothes or what happened wha- what people wear,

Speaker B cannot support speaker A's opinion on the character's clothes because she never pays attention to what people wear. Therefore, although speaker B expresses an opinion, her knowledge about the subject is, in a sense, inadequate, which is why *you know* functions as an appeal to understand the inadequacy. (Müller 2005: 182–183.)

2.2.9 Reference to shared knowledge

Reference to shared knowledge is a category where the function of *you know* comes closest to its core meaning (Müller 2005: 177). In this function, the speaker is focusing the addressee's attention to information that is shared (Müller 2005: 187; Schiffrin: 1987: 267), considered to be general information (Schiffrin 1987: 267) or regarded as common ground (Erman 1992: 222). Schiffrin (1987: 268) defines *you know* as a marker of meta-knowledge and divides it into four different types, and taking into consideration the academic setting of the ELFA corpus, types (a) and (d) seem especially relevant to this study. According to Schiffrin (1987: 268): “[i]n (a), the hearer knows the background information and the speaker knows that; ... in (d), the hearer does not know the background information and the speaker does not know that.” Therefore, for example during a lecture,

a lecturer may assume that every addressee has the same level of knowledge of the subject because it was discussed last time. The reality, however, may be that some of the addressees were not present last time but the lecturer is unaware of it.

In Müller's (2005: 178) data, the function of *you know* is to remind the addressee about a character or a scene in the movie that is important for understanding the current narrative. It was also used in argumentative discourse for similar reasons, as in example (13) (Müller 2005: 179–180):

- (13) A: the= ... the guy the thief was really convincing. .. *you know* the big guy.
 B: [yea=h].
 A: [<SV at the] beginning SV>

In example (13), both interactants know the character speaker A is talking about, and it is important to speaker A that speaker B remembers to which character he is referring. Thus, *you know* functions as an introduction to *the big guy* that serves as a reminder to speaker B.

The difference between 'reference to shared knowledge' and 'offering background information' (see 2.2.3 *Introducing an explanation*) is that background information is considered as unknown information to the addressee, which is why the speaker offers it. In 'reference to shared knowledge', however, the referred knowledge is not new to the addressee even though s/he might not immediately remember it. To know the difference between the two in the ELFA corpus lies in the context. The name of the lecture in example (14) is "QCD and Hadron Structure":

- (14) <S1> and we have these moving bound states which of course we have to have in in QCD
you know these scatter hadrons they must be moving then extremely little is done on this
 (ULEC090)

Example (14) is similar to example (13) from Müller (2005: 180): the speaker gives an argument *which of course we have to have in in QCD* and continues to explain it further in the utterance *these scatter hadrons they must be moving* after *you know*. In addition, the name of the lecture as well as

the speaker's usage of *of course* suggest that the addressees have or at least should have some knowledge of the topic of the lecture.

Being able to differentiate 'securing comprehension' and 'reference to shared knowledge' can be a little challenging as one does not know how well the interactants know each other in the ELFA corpus. Thus, in this thesis, one must rely on the transcript and the surrounding context to determine whether *you know* is used as a reference to shared knowledge or as securing comprehension. In other words, if it can be shown from the ELFA transcript that a) the speaker knows that the addressee shares the same knowledge or b) at least the speaker assumes that the addressee shares the same knowledge, *you know* is considered to be used as a reference to shared knowledge (e.g. example (14)). If, however, this is not visible from the transcript, *you know* is considered to be used as securing comprehension.

2.2.10 Acknowledge that the speaker is right

The last function of *you know* is *acknowledge that the speaker is right*. This function works on interactional level because the speaker is negotiating with the addressee to agree with the speaker (Müller 2005: 184). Müller (2005: 184–185) explains that the scale of negotiation goes from stating to asking to demanding that the listener should agree with the speaker. Stating means that the speaker is sure s/he is right or that at least s/he has the right to an opinion. Asking means that it is important to the speaker that the addressee agrees with him/her. Demanding means that the speaker and addressee have opposing opinions and the speaker is attempting to show that his/her opinions are better than the addressee's. (Müller 2005: 184–185).

Example (15) from the ELFA corpus shows a case where the speaker is stating that he is right:

- (15) <S2> i think that's maybe one of the areas where (your) way of interpreting er mouffe and er laclau is is actually better than what they do themselves *you know* (UDEF070)

In example (15), the speaker is congratulating the addressee on interpreting political theorists Mouffe and Laclau's views better than they do. The function of *you know* is similar to a tag question: the speaker is not asking or demanding the addressee to agree with him, but more likely the speaker is stating an opinion and checks that the addressee agrees with the speaker's praises.

Again, Müller (2005: 184) uses intonation contour to define the tone in which the appeal for acknowledgment is made. In this study, however, only the transcript is used to define if the speaker is making such an appeal. This is because knowing that an appeal for acknowledgment occurs at all in the ELFA corpus is sufficient for the purposes of this study.

Schiffrin (1987: 276) notes that *you know* acts as a marker of meta-knowledge, meaning that "[s]peakers often use general descriptions to support their more specific claims and to gain their hearers' endorsement of such claims." In other words, the speaker asks the addressee to acknowledge that s/he is right based on a truth that is supposed to be universally acknowledged by everyone within the same society (Schiffrin 1987: 274–276). Referring to the same passage from Schiffrin (1987: 276), Erman (2001: 1348), on the other hand, calls it the speaker's appeal "to shared knowledge of the world, general truths, or otherwise 'uncontroversial' issues". Despite the words "appeal" or "shared knowledge" that in Müller's (2005) study refer to other functions of *you know*, the meaning behind Erman's (2001: 1348) words is that generalisations are used to gain the addressees' support for the speaker's claims – thus their purpose is to ask the addressee to acknowledge that the speaker is right.

3. Corpus data and methodology

Now that the relevant terms as well as the functions of *you know* from previous literature that are relevant to this study are introduced and explained, it is apt to present the data used in this thesis, i.e.

the corpus of English as a Lingua Franca in Academic Settings, or the ELFA corpus (ELFA 2008). In addition, the method of analysis is also introduced in more detail.

3.1 ELFA

The ELFA corpus consists of one million words of transcribed spoken academic English as a lingua franca, which is around 131 hours of transcribed recordings. There are approximately 650 ELF speakers from several continents with 51 different first languages. Native speakers of English are also present in the ELFA corpus but their speech only takes 5% of the corpus. (ELFA 2008.)

The corpus consists of dialogic or polylogic speech events (67%), such as thesis defenses and seminar discussions, as well as monologic speech events (33%), such as lectures and thesis presentations. None of the speech events were elicited by the corpus compilers, but they occurred naturally, which is why the data in the ELFA corpus is considered authentic. The data was recorded at the University of Tampere, the University of Helsinki, Tampere University of Technology and Helsinki University of Technology. The disciplinary domains are also various: social sciences, technology, humanities, natural sciences, medicine, behavioural sciences and economics and administration. The project director Anna Mauranen and her team completed the ELFA corpus in 2008, but it is still being developed and is freely available for researchers. (ELFA 2008.)

The ELFA corpus consists of both transcripts and recordings (ELFA 2008). In this thesis, the transcripts were primarily used as the data, but the recordings were used whenever the transcript was insufficient in defining, e.g., if *you know* was used as a discourse marker or non-discourse marker. The abbreviations used in the ELFA transcripts are presented in Appendix 1.

Furthermore, every ELFA transcript consists of file headers that offer metadata of the transcript (ELFA 2008). In addition to the file ID, it shows the academic domain, discipline, event type (lecture/discussion/etc.) as well as possible notes, such as which other files are connected to the

transcript at hand. Moreover, file headers also include information of the participants present during the event type: the number of speakers, their mother tongues, academic roles, age, gender as well as the number of all participants.

3.2 Method of analysis

As the data for this study is from the ELFA corpus, a corpus-driven approach is utilised in this thesis; an approach, described by Tognini-Bonelli (2001: 84), that takes into account all examples of the analysed linguistic item without any modifications or limitations. Müller (2005: 12) also emphasises that in corpus-driven approach, and in descriptive language research in general, it is important to take into account all instances of the studied linguistic phenomenon and not just the most suitable ones. Therefore, all instances of the discourse marker *you know* appearing in the ELFA corpus are taken into account in the analysis.

Erman (1987: 33) distinguishes two traditions of analysing linguistic items such as *you know*: a top-down analysis and bottom-up analysis. Erman (1987: 33) explains that in the former the data is analysed through a set of pre-established functions which are desired to be discovered in the data, whereas in the latter the functions are not pre-established but arise from the data itself. The approach used in this study is more top-down than bottom-up: the data is analysed with the help of pre-established functions by Müller (2005) and House (2009) who also focus on researching non-native speakers of English. However, during the analysis of the ELFA corpus, it became clear that the functions established by Müller (2005) and House (2009) are not sufficient. For example, neither of them refer to *you know* as a marker introducing a proposition, whereas Erman (2001), who studies native English speakers, on the other hand, does. Therefore, previous studies on native English speakers are also utilised in this study to make the functions *you know* seem to have the ELFA corpus as accurate as possible.

All instances of *you know* that occur in the ELFA corpus are studied by using the AntConc programme, and with the great help of professor Sebastian Hoffmann, all instances of *you know* were sorted in detail into an excel file by using the PERL programme. The total number of instances of *you know* in the ELFA corpus is 1,272, which was received by using the search terms *you know*, *y- know*, *you kn- know* and *you kno-* on the AntConc. However, the vast majority of instances were received by the search term *you know*, as *you kn- know* occurred only once, *y- know* only twice and the only time *you kno-* occurred, it was spoken by a speaker whose second language was English, which is why it was not taken into account in the analysis.

After the initial search, all the instances of *you know* spoken by ELF speakers were divided into discourse markers and non-discourse markers. The division was based on previous studies (see 2.2 *Defining the functions of you know*) and intuition. Cases where *you know* was considered as a non-discourse marker were those in which *you know*, e.g., could not be omitted from the syntactic structure, i.e. it was not grammatically optional (Müller 2005: 5–6). This includes cases such as *do you know* questions (with or without the auxiliary verb *do*) and cases where *(you) know* takes a complement, e.g. *you know what I mean*.

After dividing the instances of *you know* into discourse markers and non-discourse markers, all instances of *you know* as a discourse marker were analysed and categorised based on their functions. The function for each *you know* was determined by comparing it to Müller (2005) and House's (2009) definitions. If, however, the pre-determined functions by Müller (2005) and House (2009) did not seem to correspond to the data, studies of native English speakers, such as Erman (1987, 1992, 2001), Schifffrin (1987), Östman (1981) and He and Lindsey (1998), were used as a reference. The ELFA recordings were only used if the transcripts were inadequate in dividing the instances of *you know* into discourse markers and non-discourse markers, as well as when the transcripts were insufficient in defining to which functional category *you know* as a discourse marker belongs. Lastly, the results

of the analysis were compared to Müller (2005) and House's (2009) studies as well as to some relevant studies on native English speakers, Erman's (2001) study in particular.

4. How is *you know* used by academic ELF speakers?

The total number of occurrences of *you know* in the ELFA corpus is 1,272 of which 1,110 is spoken by ELF speakers. ELF speakers use *you know* as a discourse marker 852 times and 247 times as a non-discourse marker. In addition, the remaining 11 times it is marked in the transcript as (*you know*), which corresponds to "uncertain transcription" (ELFA Transcription Guide 2004: 1).

As previously stated, *you know* is categorised as a non-discourse marker when it functions in its original sense and is thus excluded from the analysis. These are cases such as in example (16) where *you know* is part of a question or in examples (17) and (18) where *you know* takes a complement (see also Müller 2005: 157):

- (16) <S1> i mean what is the fundamental difference in the formation of the zambian state as compared to the formation of the zimbabwean state , do *you know* that [does can anybody comment on the] </S1> (USEMD26A)
- (17) <S9> one of them has been the choice of er selecting a husband er which is very popular in er urban areas in which girls decide to whom they gonna be married but in rural areas as *you know* the traditions still exist and er (CDIS090)
- (18) <S4> i mean a- acting on quote unquote knowledge requires that there is another higher level of organisation *you know* what i mean (USEMD130)

In addition, Müller (2005: 188) notes that in cases where the speech is unintelligible or the amount of context is insufficient, it is not possible to determine the function of *you know*. For this reason, the 11 cases that are marked unclear in the transcript are not analysed nor are they divided into discourse and non-discourse markers. The same rule applies for six of the 852 instances where *you know* is used as a discourse marker, but it is closely preceded or followed by a stretch of speech

that is marked as (xx) that corresponds to “unintelligible speech” (ELFA Transcription Guide 2004: 1). Based on the ELFA transcripts, the length of (xx) may vary from one word to multiple clauses in the ELFA corpus, which in these cases makes the amount of context insufficient. Thus, the total number of instances of *you know* in the ELFA corpus that are analysed and categorized into different functions is 846.

Before moving on to the analysis of *you know* in the ELFA corpus, figure 1 gives an overview of how frequently different functions of *you know* occur in the data:

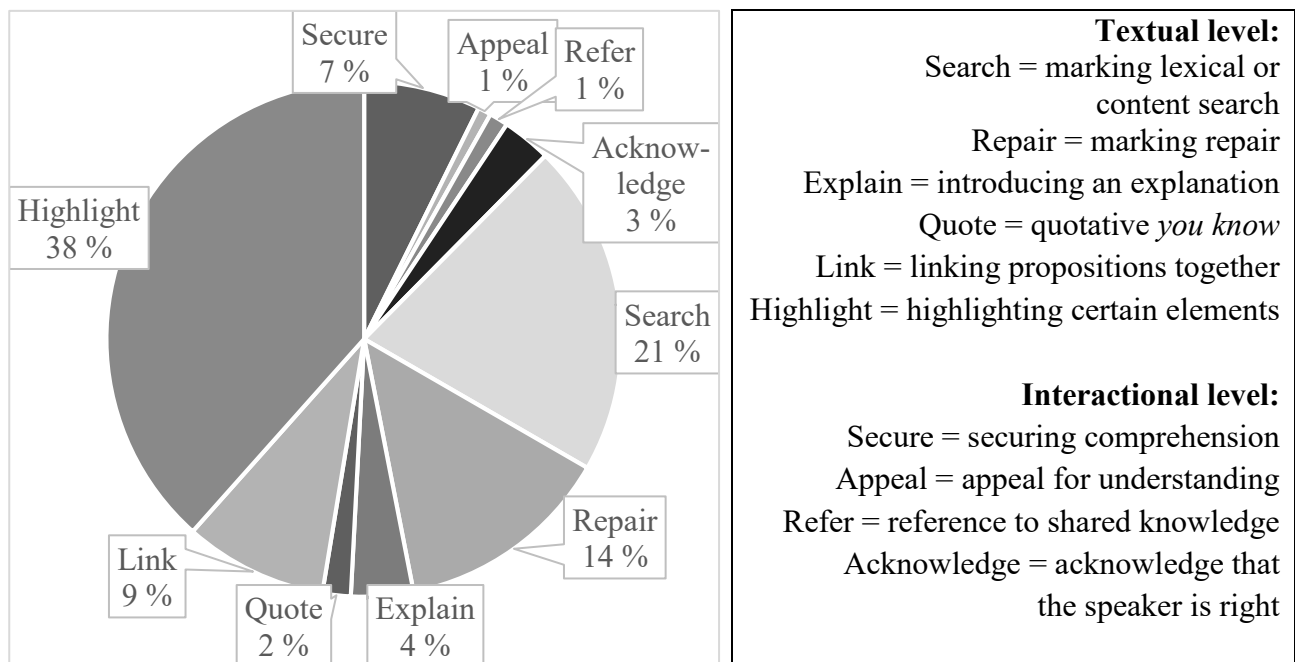


Figure 1. Overview of the frequencies of *you know* in the ELFA corpus

As can be seen from figure 1, ‘highlighting certain elements’, ‘marking lexical or content search’ and ‘marking repair’ are the three most frequent functions of *you know*, whereas ‘quotative *you know*’, ‘reference to shared knowledge’ and ‘appeal for understanding’ are the three least frequent functions of *you know* in the ELFA corpus. In the next section, every function of *you know* in the ELFA corpus is given a detailed analysis.

4.1 Functions of *you know* in the ELFA corpus

Table 2 was previously introduced in 2.2 *Defining the functions of you know* as a part of table 1. Hence, as previously mentioned, the functions shown in table 2 are partly adapted from Müller (2005). Some categories are kept partially or completely the same, e.g. ‘marking repair’ or ‘introducing an explanation’. However, two of Müller’s (2005) original functions are combined into one category, ‘securing comprehension’. Moreover, there are two completely new categories, ‘linking proposition together’ and ‘highlighting certain elements’, that are largely based on House (2009) and Erman’s (2001) studies.

Table 2. The functions of *you know* as a discourse marker in the ELFA corpus

| |
|--|
| Textual level: |
| <ul style="list-style-type: none"> - marking lexical or content search - marking repair - introducing an explanation - quotative <i>you know</i> - linking propositions together - highlighting certain elements |
| Interactional level: |
| <ul style="list-style-type: none"> - securing comprehension - appeal for understanding - reference to shared knowledge - acknowledge that the speaker is right |

It must be noted, though, that the functions above are not clear-cut but *you know* can have more than one function at a time (Erman 1987: 121). Therefore, it is the most salient function of *you know* that determines to which category *you know* belongs (see also Müller 2005: 174).

The analysis of *you know* as a discourse marker in the ELFA corpus includes the following grammatical abbreviations: NP (noun phrase), VP (verb phrase), AdjP (adjective phrase), AdvP

(adverb phrase) and PrepP (prepositional phrase). An NP consists of a noun and its possible determiner, e.g. *the, this, my*, and/or premodifiers, e.g. *happy, quite, two*. An NP can also consist of a pronoun, e.g. *I, it, we*. A VP in this thesis consists of a simple verb phrase or complex verb phrase but not its dependents, e.g. “I always *knew* that I *would become* a teacher”. An AdjP includes an adjective and its possible complement or modifier, e.g. “Your argument is *quite good*”. An AdvP consists of an adverb and possibly its modifier, e.g. “He speaks *slowly enough*”. A PrepP includes a preposition and its complement, e.g. “That depends *on how well you do*”. The abbreviations, their definitions as well as examples are compiled in table 3:

Table 3. Grammatical abbreviations used in the thesis

| Abbreviation | Definition | Example |
|--------------|---|--------------------------------|
| NP | noun (+ determiner/premodifier), pronoun | <i>(a happy) encounter, it</i> |
| VP | simple verb phrase, complex verb phrase | <i>say, should have said</i> |
| AdjP | adjective (+ complement/modifier) | <i>(really) funny</i> |
| AdvP | adverb (+ modifier) | <i>(rather) quietly</i> |
| PrepP | preposition + complement | <i>on the way</i> |

The notion of *phrase* is used in this thesis rather similarly to the notion of *group* used by Halliday and Matthiessen (2004: 54). For example in this thesis as well as in Halliday and Matthiessen (2004: 54), the notion of *NP* consists of the noun and its possible determiners or premodifiers. In addition, the notion of *VP* in both cases consists of simple or complex verb phrases but not the prepositional phrase that may follow it (see Halliday & Matthiessen 2004: 54).

4.1.1 Marking lexical or content search

You know marking lexical or content search occurs in all event types from monologues (e.g. lectures) to dialogues/polylogues (e.g. seminar discussions) in the ELFA corpus, and with 177 instances, it is one of the most frequent functions of *you know*. Müller (2005), House (2009) and Erman (2001) state that as a marker of lexical or content search, *you know* functions on textual level or micro-level, which means that *you know* is targeted towards the speaker and what is said rather than the addressee. This is also confirmed in the ELFA corpus as the target of the search is always a word, a phrase or content, i.e. what the speaker wants to say or how s/he wants to say it.

Like in Müller (2005: 158–159), the primary indicators of this function are pauses or verbal hesitation markers, such as *er* or *mhm*, which are also called filled pauses. In addition, similar to Erman (2001: 1344), repetition of function words, such as conjunctions, prepositions or articles before and/or after *you know* also indicate a search for a word or content.

The results of the analysis of the ELFA corpus support Müller's (2005: 160) suggestion of the category 'marking lexical or content search' being a continuum that consists of explicit cases of lexical search and content search, as well as cases that have elements of both in them. However, most of instances of *you know* in the ELFA corpus are targeted either towards lexical or content search, as there are 59 instances of lexical search, 109 instances of content search and nine cases where it is not certain if the speaker is searching for a word or content.

In the majority of instances where *you know* marks lexical search, it occurs in mid-utterance and the word the speaker is looking for is an NP. Usually, *you know* occurs within an NP or next to it, e.g. within a PrepP when the PrepP complements an NP or after a main verb when *you know* precedes an NP that functions as the object or complement of the verb. Occasionally, when *you know* follows a verb, the target of the search is another VP that complements the first VP. Example (19) demonstrates *you know* occurring within an NP:

- (19) <S1> i don't like them particularly but he is the name which is being talked about in photographic er *you know* circles and this now we we're going to compare you remember who is this . it is alexander rodchenko of course (USEMD060)

In the excerpt, the subject of the talk is Evgeny Mokhorev and his photographs, and *you know* appears within the NP that consists of the adjective *photographic* and the noun *circles*. There is also the hesitation marker *er* preceding *you know*. The position of *you know* indicates that the speaker is searching for a noun that *photographic* is attributing, and the function of *you know* is to stall for time as well as help the speaker to find the proper word.

When *you know* occurs within a PrepP that complements an NP, it often appears in a pattern *noun + of + noun* where the preposition *of* introduces a noun complement:

- (20) <S13> but as you said you need er another research er er set of er *you know* er interventions to to measure these er diversities in poverty in those two contexts [please] (CDIS08A)

In example (20), *you know* occurs after the first noun *set* and the preposition *of* of the pattern *noun + of + noun* and before the noun complement *interventions*, which is the searched word. In addition, the speaker seems to have difficulties in finding words or content already before *you know*. Thus, *you know* also functions as a focusing device directing the addressee's focus on the part that follows *you know* (cf. House 2009: 187), i.e. the noun complement *interventions* and what follows it.

When the target of lexical search is not a noun, it is a verb. In example (21), *you know* occurs after a VP:

- (21) <S4> after all finland is leading er information society in the world so the comparison is intended *you know* to er to to to ascertain from the facts that will be derived so far (USEMD04B)

In the excerpt, the speaker is searching for the complement of the VP *is intended*. *You know* is, however, not preceded by a pause nor is it immediately followed by one either, as there is the

preposition *to* between *you know* and the hesitation marker *er*. In fact, it is primarily the repetition of the preposition *to* that indicates a search of some sort, and based on the coherent part of speech that follows *you know*, it is most likely the verb *ascertain* that was missing than the whole content.

There are a couple of instances of *you know* marking lexical search where the immediate result is not a noun or a verb. Twice it is followed by a clause where speakers explicitly voice their difficulties in finding a correct noun or verb. In example (22), the speaker is searching for a correct noun:

- (22) <S15> i mean the summit will be er repre- representation of compromises done by you know local or i don't know regional continental *you know* er what you call them work groups so also you might find some you know clues or points or <S1> mhm </S1> you know aspects of that in the prepcom work <S9> mhm </S9> but </S15> (USEMD04B)

In the excerpt, speaker S15 explicitly comments the trouble she has in finding the right word (an NP) to which the adjectives *regional* and *continental* function as attributes. After the comment *what you call them*, S15 settles for the NP *work groups*. The other instances of *you know* in example (16) are highlighters (see 4.1.6 *Highlighting certain elements*).

Once the result of a lexical search seems to be an adjective:

- (23) <S1> so the the main idea when we designed this was that we do not make the plug-ins by ourselves to the er *you know* to the best (xx) so that we take for example M-PEG-2 decoding library and we just put that the interface to it (UDEFD080)

In example (23), *you know* is followed by *to the best (xx)*. In the ELFA corpus, unintelligible speech is transcribed as (xx) and the stretch of talk may vary from one word to multiple clauses. It is probable that in example (23), (xx) corresponds to a word, most likely an NP where *the best* functions as an attribute. Thus, it seems that the target of the lexical search in the excerpt is not the adjective *the best*, but the unintelligible word, most likely a noun, that follows it.

In cases where *you know* marks a search for content, it primarily occurs at the beginning of a clause before the main verb, i.e. in clause-initial position, either between two independent clauses or between a main clause and a subordinate clause, or vice versa. Sometimes *you know* occurs after a VP when the subject and the verb are combined, e.g. *there's*, or when the speaker forms an argument with the phrase *I think*. In these cases, *you know* is considered to occur in mid-utterance. Searching for content in all these situations is understandable because the main information of an utterance often comes after a VP. Thus, instead of searching for a particular word, the speaker is searching for content.

Example (24) demonstrates content search where *you know* functions between two independent clauses. The @-sign represents laughter.

- (24) <S4> [in] normal circumstances though i've seen humans you know wiping up other other humans you know rationalising that you know i mean which means maybe lions are a little bit more intelligent actually @@ er *you know* so all i'm saying i mean you know i mean democracy knowledge wisdom @@ i think wisdom might take us to democracy much faster you know than than than knowledge (USEMD130)

Speaker S18 is a frequent user of the discourse markers *you know* and *I mean*. In the excerpt, *you know* marking content search is preceded by an argument that ends with laughter. Therefore, *you know* begins a new clause that sums up the speaker's argument and it is preceded by a filled pause. The filled pause suggests a search for either words or content, but the position of *you know* before the VP *'m saying* points toward content search. The other instances of *you know* in example (24) are highlighters or markers of repair (see 4.1.2 *Marking repair* and 4.1.6 *Highlighting certain elements*).

When *you know* functions in close vicinity of a subordinate clause, the position of *you know* – whether it is before or after a conjunction or a relative pronoun – is not significant as both cases occur. In example (25), *you know* occurs before the relative pronoun *that*:

- (25) <S1> it has always been that they needed to be seen to be giving something in the face of their own populations perhaps demanding *you know* er that that yeah we don't just leave these people to die or fall off the planet er (USEMD26A)

You know precedes a subordinate relative clause, which is the object of the verb *demanding*. Therefore, *you know* occurs directly after the VP and is followed by a filled pause, as well as the repeated relative pronoun *that*. These, consequently, infer to content search.

There are a couple of cases in the ELFA corpus where *you know* as a marker of content search occurs in mid-utterance. In these cases, the speaker is often listing examples to support his/her argument and *you know* occurs within such a list. In example (26), speaker S2 is talking about negative connotations that the notion of national pride often has in Western Europe. S2 gives some examples:

- (26) <S2> it is it is associated more with things like er erm a kind of aggressive nationalism with war with with intolerance with erm war crimes perhaps even *you know* with with er er so obviously er relating to things like the the second world war (ULECD140)

However, the last part of the list beginning with *perhaps even with* does not receive a specific example but is substituted with a longer clause *so obviously er relating to things like the the Second World War*. The reason why *you know* in this excerpt is analysed as content search and not as lexical search is that the speaker is searching for an example, which functions on a more abstract level than searching for a specific word. For example, if a person is asked a question “What is your favourite animal?” and s/he answers “Erm, a dog”, most likely the hesitation marker is not due to the lack of a word for the aforementioned animal, but because the person has to think about the content before answering. Hence, even though the outcome of the search in example (26) would most likely have been a noun, it is not the word that is missing but the example, i.e. content. Moreover, S2 does not eventually come up with a specific example, so he substitutes it with a more generic *relating to things like the Second World War*.

Lastly, in some cases, it cannot be certainly determined whether *you know* marks lexical or content search. In these cases, the outcome of the search is often more than one word, i.e. a phrase or a metaphor. In example (27), S2 is asked a question about the relationship between Finland and the Soviet Union during the period Kekkonen acted as President of Finland:

- (27) <S5> how- was it was it actual situation or or was it just play that kekkonen played with the @with the@ soviet leaders </S5>
 <S2> well er i do believe that er the situation those days er er was *you know* was very close to explode (USEMD01A)

In the excerpt, the filled pauses before *you know* in addition to the repetition of the linking verb *was* before and after *you know* indicate that speaker S2 is searching for a way to respond. The outcome of the search is a longer stretch of talk *was very close to explode*, which is a metaphor. However, it could also have been a single word, such as *delicate*. Müller (2005: 160) points out that in some cases, it is unlikely that speakers themselves could be able to state if the reason they hesitated was due to the lack of content or because they, in fact, had the content in mind but not the words to describe it. Thus, it can be argued that the search in example (27) holds elements of both content and lexical search, which is why it cannot be placed to either ends of the continuum but somewhere in between.

To sum up the category ‘lexical or content search’ in the ELFA corpus, when *you know* marks lexical search, it often occurs in mid-utterance and the word the speaker is usually looking for is an NP. As a marker of lexical search, *you know* primarily occurs next to or within an NP or in a *noun + of + noun* pattern where the latter noun functions as a complement. Occasionally, *you know* occurs between a verb and its object or complement NP or VP. When, on the other hand, *you know* marks content search, it primarily occurs in clause-initial position either between two independent clauses or between a main clause and a subordinate clause. The usual position of *you know* as a marker of content search is at the beginning of a clause or before a VP. Sometimes *you know* also occurs right after the main VP and in rare cases in mid-utterance. When *you know* functions next to a subordinate

clause, it occurs either before or after a conjunction or a relative pronoun. Lastly, there are some cases where it is not possible to tell if the speaker is initially searching for a lexeme or content. This is because the outcome of the search is often a phrase or a metaphor.

4.1.2 Marking repair

There are 115 instances of *you know* marking repair in the ELFA corpus, which makes it also one of the most common functions of *you know* in the data. The types of repair vary from false starts and word replacements to restarts after *you know* (cf. Müller 2005, Erman 2001). Moreover, in the majority of instances, *you know* occurs in mid-utterance and in more than half of the instances after or within a VP or an NP (cf. House 2009). Lastly, as the repairs that *you know* marks deal with lexical choices, it functions at the lexical level or micro-level (cf. Müller 2005, House 2009).

Example (28) demonstrates a type of repair where nothing is repaired but repeated; a type that Müller (2005) would name as a false start. However, false starts only cover one fifth of all the repairs in the ELFA corpus, which in turn supports the decision to rename the function from ‘marking false start and repair’ to ‘marking repair’. In example (28), *you know* occurs in mid-utterance and after an NP:

- (28) <S3> so it's like the people *you know* the people they don't understand like what's happening (USEMD100)

As the excerpt shows, speaker S3 does not make an explicit repair but simply repeats the NP *the people* that also precedes *you know*.

The remainder of repairs in the ELFA corpus are either word replacements or restarts. *Word replacement* means that a complete or unfinished word is substituted with another one after *you know*, whereas *restart* means that the speaker begins an utterance but stops in the middle and begins anew

after *you know*. In example (29), *you know* marks word replacement that occurs in mid-utterance after a VP:

- (29) <S1> and er people are going to the mosque er all over the place and women have *you know* they're wearing the what is it called hajab </S1>
 <NS3> hijab </NS3>
 <S6> yeah [hijab] </S6>
 <S1> [hijab] yeah and so just just (USEMD26A)

In the excerpt, speaker S1 makes two replacements: he replaces the noun *women* with the pronoun *they* (that still refers to *women*), as well as the used main verb *have* to *wear*. In addition, S1 also makes another type of repair by changing the verb form from the simple form *have* to the progressive form *'re wearing*. The second type of repair also alters the syntactic structure of the utterance. In fact, a change in the syntactic structure is always possible when a complete or unfinished word is substituted with a verb.

Example (30) also demonstrates word replacement but this time *you know* occurs within a VP. When *you know* occurs within a VP, e.g. after an auxiliary verb and before a main verb, it is also considered to occur in mid-utterance.

- (30) <S2> normally er when we measure NGN three by a quantitative RTPCR in er in normal mice we have practically no er expression we can pra- *you know* detect it but after duct ligation this is in duct ligated pancreas part (ULEC150)

Speaker S2 begins a new clause with *we can* but cuts off in the middle of the word *pra-* and replaces it with the main verb *detect* after *you know*. Thus, S2 makes a word replacement from an unfinished word to a different yet completed one after *you know*.

In example (31), the speaker does not only replace a word but the entire utterance before *you know*. Again, *you know* occurs in mid-utterance after a VP:

- (31) <S2> finland was an ally with germany er at the end of of of world war two and even when germany was defeated there were some troops stationed at the north of finland , that has forced *you know* er from the soviet perspective they thought that finland is posing a threat because it's still keeping @german soldiers@ in the soil of finland (USEMD01A)

In the excerpt, *you know* marks a repair from one utterance *that has forced* to another *from the Soviet perspective*. In other words, the speaker “erases” the unfinished utterance said before *you know* and replaces it with a new one. This repair also helps the speaker to form his argument from another perspective.

Example (32) shows another repair where the speaker starts anew after *you know*:

- (32) <S2> then you go to usually do reductions that are going to make it impossible to represent this , so maybe you need to to er *you know* when you're doing your reductions keep more state information to make it possible to (UDEFD060)

In example (32), the function of *you know* is similar to the function of ‘marking lexical or content search’: S2 begins his suggestive argument with *so maybe you need to* but cuts off after it and despite the repeated preposition *to* and the hesitation marker *er* is unable to finish the argument. Thus, S2 begins a completely new utterance after *you know*, i.e. restarts with a changed syntactic structure. The reason why *you know* is regarded as marking repair and not as marking lexical or content search in example (32) is that the function of repair is more dominant than the function of lexical or content search. If the outcome of the search had been a successful lexical or content search, it would have been regarded as a marker of lexical or content search.

Although most of the repairs occur in mid-utterance in the ELFA corpus, there are some occasions where *you know* occurs at the beginning of an utterance and before a VP, i.e. in a clause-initial position. Such a case is shown in example (33) where *you know* occurs after the subject NP but before a potential VP that is left unsaid due to the repair:

- (33) <S17> another interesting finding is that kiosk customers are almo- er almost always the same sex as the person running the kiosk and so er and we *you know* what we found through asking erm additional questions is that (CPRE08F)

The speaker replaces the previous utterance *and we* with *what we*. In other words, the conjunction *and* is replaced by an interrogative word *what*, which also alters the syntactic structure of the previous utterance.

In conclusion, as *you know* marking repair deals with lexical choices, it functions on textual level. The majority of repairs in the ELFA corpus are either word replacements or restarts, whereas the remainder consists of false starts. Sometimes a word replacement may also lead to a change in the syntactic structure. The syntactic structure of an utterance also changes when the speaker comes to a halt and starts anew after *you know*. The position of *you know* as a marker of repair is usually in mid-utterance and after or within a VP or an NP.

4.1.3 Introducing an explanation

There are 33 instances of *you know* introducing an explanation in the ELFA corpus, which means that it is one of the less common functions in the ELFA corpus. Based on the data, *you know* primarily introduces an explanation to what is said before *you know* by modifying it, rephrasing it or introducing a parenthetic comment that offers additional information for the addressee to understand the speaker's narrative. Thus, the reason for using *you know* to introduce an explanation is to avoid or prevent misunderstanding or misinterpretation.

In the ELFA corpus, *you know* introducing an explanation occurs primarily in discussions that in an academic setting often consists of argumentative discourse. Hence, discussions and argumentative discourse seem like a natural environment for *you know* introducing an explanation to occur as discussions are unplanned speech events (compared to, e.g., lectures), yet speakers aspire to explicit and clear argumentation.

In the ELFA corpus, *you know* often introduces an explanation that modifies the previous utterance and, usually, the modification makes the previous utterance more specific. Specification may span from an added word to a clause like in examples (34) and (35), respectively:

- (34) <S34> it's a question of er market is *you know* a technical question er if the market was expanded if there was some kind of de-regulation (CDIS08B)
- (35) <S17> they make sure they point at it and they get it *you know* someone is going to get it for them if they want or if they don't want something (CDIS06B)

In example (34), the speaker adds the adjective *technical* to specify the noun *question*, an addition that the speaker most likely considers relevant for the addressee to know in order to understand the speaker's argument. In example (35), on the other hand, the speaker is talking about babies and their ability to get what they want. However, knowing that babies who are unable to walk cannot get objects themselves, the speaker modifies the clause *and they get it* before *you know* into a more specific *someone is going to get it for them*.

Sometimes the explanation *you know* introduces is a rephrased utterance, which means that the part said before *you know* is said again after *you know* but in different words. Usually the angle changes a little as well, like in example (36):

- (36) <S15> the way in which we refer to to people already gives us er an an impression of what our position is er with respect to the person or the or the or the or the thing modality er both er epistemic and deontic *you know* the level of certitude we we we assign to what we say (CPRE06C)

The repetition of *or the*, the hesitation markers as well as the word *the thing* indicate that speaker S15 is having trouble finding the proper words before *you know*. It seems, though, that after *the thing*, she finds the correct word *modality*, but explains it further with the clarifying concepts *epistemic* and *deontic*. However, after *you know*, the speaker says the same idea of epistemic and deontic modalities but with different words and from a slightly different point of view, saying *the level of certitude we*

we we assign to what we say. The part following *you know* also seems to sum up what the speaker means with epistemic and deontic modalities and how they are related to her argument.

Like modification, parenthetic comments are also common in the ELFA corpus when *you know* introduces an explanation. The function of a parenthetic comment is to offer additional information for the addressees that helps them understand what the speaker is saying. After the added parenthetic comment, the speaker continues the previous utterance, often repeating a part of what occurred before *you know*. This is the case in examples (37) and (38):

- (37) <S17> and what that means is that the kiosk operator is basically mediating the interaction between the PC and the person so especially for us we're er *you know* as a research lab we're ultimately interested in finding technical innovations that might help these rural projects (CPRE08F)
- (38) <S2> but when when she was <COUGH> translated into french , *you know* in the swedish original version pippi longstocking she's extremely strong and she's lifting er er her own horse like <S1> [mhm mhm] </S1> [it's it's no] no problem for her but in the french version she only lifts a pony <SS> @@ </SS> (USEMD150)

A parenthetic comment may be short, as in example (37), or a longer stretch of talk, as in example (38). In example (37), speaker S17 begins a new clause with *we're* but decides that more background information is needed. Therefore, S17 adds *as a research lab* and then continues the previous utterance by repeating *we're*. In example (38), speaker S2 is comparing different translations of Pippi Longstocking and begins a new clause with *when she was translated into French*. However, in order for the comparison to be explicit to the addressees, S2 gives additional background information about the original version of Pippi Longstocking that the speaker considers important for the addressees to know. After the parenthetic comment, the speaker continues the previous utterance and repeats the word *French*, which shows that the parenthetic comment is ended.

In conclusion, *you know* often introduces explanations that modify the previous discourse or introduces a parenthetic comment that offers additional information for the addressee to understand the speaker's narrative in the ELFA corpus. In addition, sometimes the speaker offers an explanation

by rephrasing what was said before *you know*. As a marker of introducing explanation, *you know* occurs primarily in lecture and seminar discussions as well as in argumentative discourse, and its purpose is to introduce an explanation that helps the addressee to understand the speaker's argument.

4.1.4 Quotative *you know*

There are 15 instances of quotative *you know* in the ELFA corpus, which makes it one of the least common functions. In the ELFA corpus, quotative *you know* occurs both in monologues (i.e. presentations) as well as in dialogues/polylogues (i.e. discussions). The position of *you know* is before reported speech and usually it introduces direct-speech reports. Moreover, in the majority of instances, *you know* co-occurs with an explicit quotative verb (see Müller 2005: 169).

The most common quotative verb co-occurring with *you know* is *SAY* in its various forms:

- (39) <S17> in some cases these kiosks were closed if the person decides to make a business decision simply saying *you know* that i am paying this much for let's say connectivity and maintenance of the PC and that's nothing in accrue nothing for my making of the kiosk erm it's er this is i think one of the more erm difficult findings for this area (CPRE08F)

In example (39), *you know* occurs in between the quotative verb *saying* and the relative pronoun *that* that together with *you know* lead to the direct-speech report. The speech report ends with the hesitation marker *erm*. In the excerpt, S17 is referring to a hypothetical speaker that serves as a rhetorical device. In other words, the reported words are not “real” but hypothetical to demonstrate speaker S17's point.

There are three occasions where *you know* does not occur with the verb *SAY*. Once, *you know* co-occurs with the verb *THINK*, which is, according to Müller (2005: 169), also a common quotative verb. Once, however, *you know* occurs on its own without any explicit quotative verb, but it is partnered with the relative pronoun *that*. In example (40), the subject of talk is the Association of Southeast Asian Nations (ASEAN) and if the association is able to create a sense of community to all Southeast Asians:

- (40) <S6> and i guess in that way there can be a sort of feeling of of community through through this organisation i don't know about the people if they know there in the countryside that *you know* yes we are , @belong to the ASEAN@ but at least they identify themself as the south-east asian people , (USEMD110)

Speaker S6 gives a direct-speech report that originates from a hypothetical group of people from the Southeast Asian countryside. The speech report is introduced by the relative pronoun *that* as well as *you know* that follows it.

In the third occasion when *you know* does not co-occur with the verb *SAY*, it co-occurs with the construction *BE + like*. In example (41), the discussion is about democratisation in the third world countries such as Tanzania:

- (41) <S3> it doesn't seem like immediate problem when you don't know why what you're going to eat or when you don't know like if you'll be able to er to buy medicine if you are very sick because there's no medicine in hospital then the whole democracy thing is like *you know* let's just leave it to the politician people they don't er able to identify themselves with democracy </S3> (USEMD100)

Again, the origin of S3's direct-speech report is from a hypothetical group of people living in the rural areas and it is introduced by *is like you know*. Interestingly, this is the only instance where the construction *BE + like + you know* is used in the entire ELFA corpus.

As stated above, in the majority of instances *you know* introduces a direct-speech report in the ELFA corpus. There are, however, a few cases in which the report is indirect, i.e. paraphrased. Such a case is in example (42):

- (42) <S12> i think <NAME> was in this one meeting where the this one old engineer , had a comment that he said *you know* , every time he flies the plane he rather flies the plane which is designed by an engineer than a sociologist <SS> @@ </SS> (CPRE03D)

In the excerpt, S12 begins the reported speech with the most common quotative verb *said* but instead of changing the deictic centre (cf. Müller 2005: 169) to the first person singular (e.g. “every time *I* fly

the plane”), speaker S12 keeps the deictic centre of the third person singular and, thus, paraphrases what the other person has said.

To conclude, ‘quotative *you know*’ is one of the least common functions in the ELFA corpus. *You know* always occurs before reported speech and usually the speech report it introduces is direct. In addition, *you know* often co-occurs with an explicit quotative verb that is usually the verb *SAY* in its various forms.

4.1.5 Linking propositions together

‘Linking propositions together’ is a function that occurs in the ELFA corpus 76 times. In most of the instances, *you know* occurs in discussions but sometimes it also occurs in presentations or lectures. Moreover, as the corpus consists of academic discourse, it is understandable that *you know* usually occurs in argumentative discourse, linking together different parts of the argument. The position of *you know* is often between independent clauses.

In example (43), speaker S2 is giving an argument that is based on studies by other scholars:

- (43) <S2> <SIGH> yes yes and , yes and and so so now that i said this i also realise that is er it has also been argued by scholars that actually american identity is not so ethnically and culturally neutral *you know* it is of course it's resting strongly on the english language and also it's a in some sense quite strongly on WASP identity (ULECD140)

The position of *you know* in example (43) is common in the ELFA corpus: first, the speaker gives a proposition *actually American identity is not so ethnically and culturally neutral* and offers proof in the following utterance after *you know*. Thus, the function of *you know* is to mark the transition between the two as well as link them together. Moreover, in example (43) and in many other similar instances, *you know* could be viewed as a relational phrase that functions like, e.g., the conjunction *because* (see also House 2009: 184).

However, finding a conjunction that *you know* corresponds to in the ELFA corpus is not always explicit nor does it seem necessary. In example (44), speaker S5 is explaining that in Japan, people are not allowed to talk on their mobile phones while using public transportation and that in turn prevented her from getting in touch with her friends. The first *you know* is linking propositions together while the latter one functions as a quotative *you know* (see 4.1.4 *Quotative you know*):

- (44) <S5> but in japan it's not possible and then i couldn't get contact with my friends so often *you know* we we called each other so many times and i got a a i got a phone call from a certain person and then i tried to call him and it's again it's somebody else saying you know that person is er er couldn't answer (USEMD190)

The function of *you know* in example (44) is to make the relation between the two independent propositions more salient. In other words, *you know* functions as a cohesive device. Unlike in example (43), however, *you know* does not seem to function instead of a conjunction but on its own, showing that the clauses surrounding *you know* are closely connected although not syntactically dependent. Example (44) also demonstrates one of the less common instances of *you know* occurring in narrative discourse and marking a transition between states in the narrative (see also Erman 2001: 1343).

Similar to example (44), it is questionable what kind of cohesive device *you know* is replacing in example (45). This time the excerpt is from argumentative discourse:

- (45) <NS3> you raised a good question about how how to er how exactly it has to be implemented which is also <S1> mhm </S1> problematic </NS3>
<S1> it is of course a very interesting question but i think that your case actually illustrates that er in some cases at least conditionality also can can go both ways *you know* the conditionality involved here is not just that if you don't have good governance we won't give you money but okay (USEMD26A)

Example (45) is a schoolbook example of Erman's (2001: 1343) definition of *you know* linking propositions together occurring in argumentative discourse and in between "the speaker's position and the backing up of it". Speaker S1 introduces a proposition *but I think ... conditionality also can can go both ways* and continues after *you know* with a supportive argument to the previous

proposition. However, it is more difficult to define if *you know* is acting instead of, for example, a conjunction (e.g. *so*), an idiom (e.g. *that is*) or a punctuation mark (e.g. a full stop). This, consequently, suggests that such a definition is not necessary to make, but *you know* could be viewed as having similar function as the aforementioned cohesive devices.

Occasionally, *you know* occurs in between a proposition and additional elaboration. However, unlike in *you know* introducing an explanation, the elaboration in example (46) does not seem to be due to the fear of being misunderstood or being unclear:

- (46) <S11> i developed some kind of new methods @when@ i was read this article i just thought that it's we- i every every time i when i read a sentence i forget about it <SS> [@@] </SS> [wh- what is it all about] so i just started er wri- wri- writing the summary when i was reading *you know* if i found something interesting i write down in the summary and then it like the summary was from some v- very difficult different er thoughts and then i tried to relate them to the structure of this and i don't know it was kind of interesting new method </S11> (USEMD220)

Example (46) is a part of speaker S11's narrative, where she explains her newly learned study method. *You know* occurs between two similar states of the narrative: the first state *so I just started er wri- wri- writing the summary when I was reading* is elaborated by an example *if I found something interesting I write down in the summary*. It does not seem, though, that the addressee's understanding of the narrative would be compromised had the speaker not added the second state of the narrative. In fact, it seems that the added example is given for the sake of the narrative. Moreover, if *you know* is viewed from House's (2009) point of view as a replacement of a cohesive device, such as conjunctions, *you know* could be seen as replacing the conjunction *and*, but it could also be replacing a colon or a linking word such as *therefore* as well in example (46).

In conclusion, *you know* linking propositions together usually occurs in discussions and during argumentative discourse in the ELFA corpus. The position of *you know* is often between independent clauses and its function is similar to cohesive devices, such as conjunctions or punctuation marks in

written text: to create cohesion and aid the addressee in interpreting the message. Hence, *you know* functions on textual level.

4.1.6 Highlighting certain elements

You know highlighting certain elements is clearly the most common function of *you know* in the ELFA corpus with 325 instances. The purpose of *you know* in this function is to highlight and guide the addressee's attention to the important part of the utterance said after *you know*, which is why it functions primarily on textual level. The position of *you know* varies: it may be after or within VPs or NPs, before or after conjunctions or the relative pronouns *that* and *which*, or after AdvPs or PrepPs that may also function as adverbial phrases. It may also be between a subordinate clause and a main clause or at the beginning of an utterance if the speaker is trying to take the next turn. *You know* highlighting certain elements is used primarily in argumentative discourse and in most of the cases in the ELFA corpus, it occurs during discussions.

The element *you know* highlights is either a proposition that follows *you know* or a conjunction that precedes or follows *you know*. However, when *you know* co-occurs with conjunctions, the conjunction often also leads to a proposition. Thus, similarly to *you know* linking propositions together, *you know* highlighting certain elements is connected to propositions as well. The main difference that sets the two functions apart is that when *you know* links propositions together, the propositions must be – as Erman (2001: 1344) expresses – “largely independent”, i.e. usually between two clauses. However, when *you know* highlights certain elements, it often occurs within a clause, i.e. during one syntactic structure. There are some exceptions to the rule, however, and they are explained later in this subchapter.

As previously mentioned, the ELFA corpus consists of monologues, such as lectures or dissertation presentations, as well as dialogues/polylogues, such as seminar discussions or doctoral

defence discussions. In all of these different types of academic discourse, propositions (e.g. arguments, claims, suggestions or questions) are a highly significant trait. Therefore, as *you know* often highlights a proposition, it is easy to understand why *you know* is the most common function in the ELFA corpus.

In example (47), *you know* occurs within a VP, and the proposition it highlights is a suggestion speaker S12 makes to the addressees in general as well as to speaker S19 in particular:

- (47) <S12> thank you <NAME S19> <APPLAUSE> i think as it's already quarter past unless there's something really urgent @@ @i i could@ call it a day and we could *you know* ask ask <NAME S19> and comment <NAME S19> afterwards like stay here to talk with her if that's possible <S19> mhm </S19> (CDIS06B)

There are two proposition in example (47) and *you know* occurs in the latter where S12 suggests that those who wish to ask questions or comment S19's presentation can do so after the conference. The proposition begins with *and we could*, and *you know* highlights the proposition that is a typical example of a longer stretch of talk, ending to the response *mhm* speaker S19 gives.

In example (48), on the other hand, there are two instances of *you know*, in which the first comes after a VP and the second after a conjunction. The first *you know* highlights the speaker's proposition, which is in the form of a question, and the second *you know* highlights the preceding conjunction:

- (48) <S2> i guess the dots mean it carries on its infi- infinite states er <S1> yes </S1> now my question is *you know* do you really need this because *you know* later you talk about büchi automata (UDEFD060)

The clause *now my question is* already guides the addressee attention to the upcoming question and thus the function of *you know* is to highlight the question even more. The second *you know*, however, emphasises the causal relation *because* creates between the surrounding clauses.

In example (49), *you know* occurs within an NP, primarily highlighting the latter part of the NP:

- (49) <S14> erm i i would actually like to re-focus the di- discussion a little bit and push it push you further er about the question of er HIV AIDS <SS> [mhm] </SS> [i mean] we have here a country with a population of 1.7 million and the the er infection rate in adult population is 38 per cent <SS> mhm-hm </SS> i mean you can't pass that fact how does it affect the national economy it has very clear consequences to life-expectancy it has very clear @consequences@ <SS> [mhm-hm] </SS> [to er dependen-] dependency ratios i mean where how does it claim to this picture we can't just just pass it and talk <S6> [erm] </S6> [about these] *you know* <S6> yeah </S6> macro-level economic [processes] <S6> [mhm-hm] mhm-hm </S6> i mean where is it </S14> (USEMD26B)

Speaker S14's long proposition begins already from *I mean we have a country with a population of 1.7 million* and *you know* occurs at the end, highlighting the NP *macro-level economic processes*.

Like in examples (47) and (48), the sentence *we can't just just pass it and talk about these* before *you know* begin to guide the addressee's attention to the following sequence and the function of *you know* is to highlight the important part of the proposition. Hence, it is also important to note that although *you know* particularly highlights what is said after *you know*, the utterance before *you know* also has a vital role in guiding the addressee's attention.

Another common place for *you know* to occur as a highlighter is before or after conjunctions. When *you know* occurs after a conjunction, it is clear that it occurs within a clause, and its purpose is to highlight the function of the preceding conjunction as well as the part following *you know*. In example (50), speaker S15 is discussing Sami people. *You know* occurs after the conjunction *but*:

- (50) <S15> they feel that nobody shows any interest for their culture and they have been marginalised to a certain extent to the point that they well not that they are ashamed of themselves but *you know* they really don't vocalise themselves to the extent that they might deserve (USEMD04A)

In the excerpt, *you know* occurs within the subordinate clause beginning with *but* and its purpose is to highlight the proposition following *you know* as well as the adversative link *but* creates between the two propositions surrounding the conjunction and *you know* (see also the latter *you know* in example (48)).

However, when *you know* occurs before a conjunction, it is less clear whether *you know* is a part of the clause before the conjunction or if it begins the clause after it. These cases are one of the exceptions to the rule mentioned above, cases where *you know* occurs between clauses but is regarded as a highlighter nevertheless:

- (51) <S3> i simply thought that this quotation is a very good example of <S2> yeah </S2> erm erm of an analysis er that can be cross-cultural in so <S2> [mhm] </S2> [far as] patriarchy as a category can work also in er this er erm let's say post-colonial context <S2> mhm </S2> because exactly of these er <S2> [yes] </S2> [power] relations <S2> [mhm-hm mhm-hm] </S2> [*you know*] and i think it is still a very good er at least from my experience er of er , fighting between er er two different models of woman (UDEFD110)

Based on the transcript, one cannot be certain if *you know* ends the previous clause *because exactly of these er power relations you know* or if it begins the next clause *you know and I think it is still a very good er... models of woman*. Thus, it occurs between the two clauses, i.e. between two largely independent propositions. The reason why example (51) and other similar instances in the ELFA corpus are not defined as *you know* linking proposition together but as *you know* highlighting certain elements is due to the conjunctions that link the surrounding clauses together. If *you know* is left out from examples (51), the link between the propositions would still exist due to the conjunction *and*. The same rule applies to example (50) as well. Thus, the overall function of *you know* – whether it occurs before or after a conjunction – is to emphasise the relation the conjunction makes as well as to highlight the proposition after *you know*.

The instances in which *you know* occurs before a conjunction are approximately a third of the instances where *you know* co-occurs with a conjunction. Thus, approximately two thirds of the instances of *you know* occur after a conjunction, i.e. within a clause. Moreover, in cases where *you know* is considered to highlight a conjunction, the distance between conjunctions and *you know* is relatively short in the ELFA corpus, *you know* occurring often immediately next to it. Furthermore,

when *you know* does not occur immediately before or after a conjunction, the distance is not more than a word or a phrase:

- (52) <S17> what kind of spending habits they have do they have the kind of er disposable income that might allow them to spend money on a rural kiosk and also *you know* how what's their er typical usage for for the kiosk , (CPREF08F)

In example (52), the conjunction *and* and *you know* are separated by the adverb *also*. Both *and* and *also* connect the two propositions together and the purpose of *you know* is to emphasise that connection as well as to highlight the second proposition following *you know*.

Instances where *you know* follows or precedes the relative pronouns *that* and *which* are similar to instances with conjunctions: in addition to highlighting the part following *you know*, the purpose of *you know* is also to highlight the link these relative pronouns create between clauses. *You know* occurs much more frequently with *that* than with *which* in the ELFA corpus, and interestingly, *you know* usually occurs after *that* but before *which*. Examples (53) and (54) demonstrate both cases, respectively:

- (53) <S3> thank you have you any other comment to make on the </S3>
 <S2> mhm i i i just would like to say that *you know* even though erm we haven't still explored all the all the potentialities so to speak of this framework i like a lot the idea underneath erm this CF-4 <S3> okay </S3> (UDEFD120)
- (54) <S1> and you are trying to pro- show your point then you are faced with this kind of a <S2> mhm </S2> erm argument this kind of er everyday thinking *you know* which is based on these er unquestioning er attitudes towards you know our nation our national history so (USEMD280)

Similar to conjunctions, when *you know* comes after a relative clause, it clearly belongs within a clause, like in example (53), whereas if it occurs before a relative clause, the case is not as clear, like in example (54). However, like with conjunctions, the absence of *you know* would not remove the

link that is already there due to the relative pronouns, which is why these and similar instances are defined as *you know* highlighting certain elements and not as *you know* linking propositions together.

Sometimes *you know* also comes after AdvPs or PrepPs that may also function as fronted adverbial phrases (cf. Erman 1987: 130; 2001: 1343) in the ELFA corpus:

- (55) <S2> it it it preserves information about the branching structure <S1> yes </S1> that i agree that's why i say it's not possible to add the weakest equivalence in this notion but but i'm trying to get a er i'm trying to get a feeling of what exactly *you know* you need to add to to observation equivalence in order to preserve livelocks . (UDEFD060)

In example (55), the speaker's proposition begins with *but but I'm trying to get...* and *you know* occurs after the adverb *exactly*. Like in examples (47)–(49), the clause before *you know* functions as a prelude and the purpose of *you know* is to highlight the important part of the proposition *you need to add to to observation equivalence in order to preserve livelocks*.

In example (56), *you know* is preceded by a PrepP that can also be called a fronted adverbial:

- (56) <S13> yeah i wanted to say the same thing like i i went to turku [a few days ago] </S13>
 <S14> [(i went there too)] </S14>
 <S13> @you too@ <SS> @@ </SS> and in the train *you know* erm when you want want to go to the toilet they have like this sign and it's red when it's when there's someone using the toilet <S23> mhm </S23> and it's green when you can go and they have the same thing for the mobile phone (USEMD190)

The proposition in example (56) begins with *and in the train*, where the PrepP *in the train* functions as a fronted adverbial. Example (56) also demonstrates the minority of instances where *you know* does not occur in argumentative discourse but in narrative discourse. Thus, the proposition *you know* highlights is part of the narrative, and the fronted adverbial sets the scene for the narrative. This is what Erman (1987: 130–131) calls *you know* organising the thematic structure of the sentence as well as *you know* introducing an argument.

Instances where *you know* occurs between clauses are also exceptions to the previously mentioned rule. However, these instances are rare in the ELFA corpus:

- (57) <S2> i've worked a fair amount on temporal logic er in my research and when you're dealing with temporal logic *you know* fairness is not really a problem because you can express it quite naturally in the language , (UDEFD060)

When *you know* highlighting certain elements occurs between clauses in the ELFA corpus, the preceding clause is usually a subordinate clause and the following clause is a main clause, like in example (57). Thus, *you know* highlights the main clause that covers the important part of the proposition, whereas the preceding subordinate clause functions as a prelude to the proposition.

Sometimes *you know* occurs at the beginning of a clause. These are instances where the speaker takes or is trying to take the turn. Frequently, turn-taking is initiated by the speaker him/herself but sometimes it is done by the previous speaker. In example (58), the speaker himself takes the turn:

- (58) <S2> yeah it seems er yes er i'm very sorry it seems that they mis- they , did a misprint here
</S2>
<S3> yeah </S3>
<S2> [because obviously] </S2>
<S3> [*you know*] when i was reading it i i got here very mixed up so i i thought that i understood something and and then then they were all all upside down because (UDEFD130)

In addition to turn-taking, *you know* once again functions to highlight the proposition following *you know*.

To summarise, the function of *you know* highlighting certain elements is the most common function of *you know* in the ELFA corpus, and as the name suggests, its purpose is to highlight as well as to guide the addressee's attention to the important part of the proposition said after *you know*. *You know* highlighting certain elements is used primarily in argumentative discourse as well as during discussions. *You know* often occurs after or within VPs or NPs, as well as before or after conjunctions.

Sometimes it occurs next to the relative pronouns *that* and *which*, after AdvPs or PrepPs or at the beginning of a clause. *You know* may also occur between a subordinate clause and a main clause but these cases are not common in the ELFA corpus. When *you know* co-occurs with a conjunction, its function is to emphasise the relation the conjunction makes as well as to highlight the part of the proposition after *you know*. Furthermore, in all cases where *you know* functions as a highlighter, the utterance before *you know* also has a vital role in guiding the addressee's attention to the important part of the proposition (except for *you know* occurring at the beginning of a clause). When *you know* occurs at the beginning of a clause, it functions as a turn-taking device as well as a highlighter to the following proposition. Moreover, although *you know* highlighting certain elements functions primarily on textual level, it also helps the addressee to understand what is the important part of the speaker's utterance and thus to make correct interpretations of the speaker's message. Therefore, *you know* highlighting certain elements can also be regarded to partially function on interactional level, which is the level where the following functions operate.

4.1.7 Securing comprehension

There are 62 instances of *you know* securing comprehension and it is the first function that primarily functions on interactional level in the ELFA corpus. The purpose of *you know* securing comprehension is to check that the addressee is able to understand the speaker's previous utterance or the implication that the speaker is trying to make with the previous utterance. Thus, *you know* is used to initiate the addressee's involvement. The most common position of *you know* in this function is clause-final where it can be compared to a tag question (see also Crystal 1988: 47). As there is only one speaker who uses *you know* to secure comprehension during a lecture (see, e.g., example (10)), *you know* securing comprehension occurs primarily during discussions as well as in argumentative discourse.

Because nonverbal activities, such as eye contact or nodding, are not marked in the ELFA transcripts, the only cues to exhibit that *you know* securing comprehension has successfully fulfilled its function are explicit verbal responses, such as back-channelling or laughter. However, 71% of the instances of *you know* do not receive a response, whereas with the remaining 29% of instances, the response varies from back-channelling to turn-taking. In cases where *you know* receives a response that turns into turn-taking, the new speaker is mostly selected by a) the current speaker or b) the current speaker or the new speaker him/herself. The optionality in b) is due to the lack of non-verbal information (see, e.g., example (60)). With the remaining instances, *you know* either receives a minimal response or the new speaker is self-selected.

Turn-taking is the most common type of response in this function in the ELFA corpus. In example (59), speaker S1 asks a question from S6, i.e. selects S6 as the next speaker. *You know* functions to check that S6 has understood the meaning of S1's question:

- (59) <S7> for example this simple example was the opening of the door and who should go first and then a radical feminist would say you should go first </S7>
 <S4> yeah <SU> [yeah] </SU> [and start] a discussion [about it @@] </S4>
 <SS> [@@] </SS>
 <S1> alright well mhm well do you think in your case that , it was done normally these things that you were put into in your place , do you think they did this i- knowingly *you know* [would you] </S1>
 <S6> [erm mhm] no i don't think so er i don't know it was only er their their habit (xx) their mhm invisible habit maybe i or maybe in their position <S1> yeah </S1> as i know it <S1> mhm </S1> mhm </S6> (USEMD050)

In example (59), speaker S1 directs her question to S6, and *you know* elicits a response from S6, which in turn results S1 and S6 to overlap. One cannot be certain, however, if the reason for overlapping is because a) S1 is not willing to yield – at least at that point – the turn to S6 after *you know* or b) S6 gives a non-verbal cue that S1 understands as a hint to, e.g., clarify her question. Nevertheless, the overlapping results to S6 taking the floor in the end.

Often, like in example (59), the current speaker selects the next speaker by asking him/her a question. In addition to question–answer pairs, other adjacency pairs in the ELFA corpus are, e.g., argument–comment pairs or request–acceptance/rejection pairs. Sometimes, however, knowing the adjacency pair or the context is not sufficient in defining who does the selection, the current speaker or the next speaker him/herself. Example (60) is a part of a seminar discussion, and all speakers except for S7 are master’s students. The discussion is about the Emerald Network and an essay S8 has previously written about it:

- (60) <S8> [no yeah] the emerald network was the network which was europe <S7> [yeah] </S7>
 [and] also er countries to the south of europe like er i think some north african countries [as
 well (xx)] </S8>
 <S3> [oh] </S3>
 <NS4> [oh okay] </NS4>
 <NS2> [yeah and yeah it's] so that the other (xx) </NS2>
 <S3> @okay@ so emerald network we have to <NS4> [yeah] </NS4> [look at] that as
 <NS2> [yeah] </NS2> [well] so we have we can [i think we can make a] </S3>
 <S8> [maybe i still have] that from my </S8>
 <NS2> do you [have still the (xx)] </NS2>
 <S3> [yeah oh] </S3>
 <S8> [yeah i still have like the references] of it and </S8>
 <SS> mhm okay </SS>
 <NS2> [we could look at the (xx)] </NS2>
 <S8> [and a short explanation if you want *you know*] </S8>
 <NS4> yeah i think that [would be] </NS4>
 <S3> [yeah that's] okay yeah <SS> yeah mhm </SS> yeah i think we can make a very b-
 big thing about er legal @protection@ </S3> (USEMD200)

As the excerpt above shows, the conversation consists of many speakers who constantly overlap. The essay S8 has written is of interest to many people (S3, NS2 and NS4), because it contains information that seems relevant to their current studies. The prelude to the offer–acceptance pair, where S8’s implied proposition is the offer, begins already with S3 stating the need for doing more research on the Emerald Network, @okay@ ... *we can make a*]. S8 then voices the possibility of still having the essay regarding the Emerald Network, which leads to NS2 to respond with a question, *do you [have still the (xx)]*. S8 answers NS2’s question affirmatively, *[yeah i still have like the references] of it*

and [and a short explanation, and begins an offer, *if you want you know*]. However, she does not finish the offer but uses *you know* to imply that she would try to find the essay and give it to the others if need be. That, in turn, leads to NS4's acceptance. S3 might also be accepting S8's implied offer with *yeah that's okay yeah* although her and NS4's turns do not immediately overlap. Therefore, without any visual information, one cannot be certain if S8 directs the offer and thus yields the turn to one certain addressee or if it is directed to everyone involved in the conversation. However, whether *you know* is directed to *you* singular or *you* plural, the main point is that *you know* is used to invite the addressee(s) to take part in the conversation, thus functioning on interactional level.

There is also one instance of *you know* securing comprehension where the speaker directs *you know* to a certain addressee but does not receive a response despite his request. In example (61), speaker S1 has asked a question from the addressees before the excerpt. The responses he receives, though, are not what he would have preferred:

- (61) <S1> [yeah yeah mhm yeah] . any other peop- please , yes japan please go ahead @@ sorry i forgot your name yeah </S1>
 <S9> er i don't know (xx) </S9>
 <S1> yeah that's alright yes any other . please finland . just say something *you know* erm
 <COUGH> well as the matter of fact i i i agree with (ULECD030)

After receiving a response, *I don't know*, from speaker S9, S1 attempts to allocate the turn to "Finland". The longer pause after *please Finland*, the encouraging implication *just say something you know*, as well as S1 continuing his turn after *you know* indicate that the response S1 receives is not verbal yet the message is the same as with S9 – "Finland" rejects S1's request. Therefore, although S1 tries to imply that any answer contributing to the question at hand is a welcomed one, he does not receive it.

As mentioned above, in most of the cases *you know* does not receive a verbal response. Such a case is presented in example (62) where the latter *you know* functions like a tag question:

- (62) <S1> if childhood is just discourse cultural construction what do you do with the children that sit in in the benches at school what do you do with children you know in that have emotions that say things that we see that are flesh and bone *you know* and this is er one of the reasons why i wouldn't use a notion as er in childhood representation research (UOTH010)

There are two instances of *you know* in the excerpt above. The first is a highlighter, whereas the second seems to be oriented towards the previous utterance underlining what is meant by it despite the lack of verbal response. However, the second *you know* might also be highlighting the conjunction *and* that occurs after it. Therefore, the recordings of the ELFA corpus were used to check if *you know* had a rising, even or falling intonation. In example (62), the latter *you know* has a rising intonation that corresponds to *you know* used like a tag question (Crystal 1988: 47), which in turn leads it to be analysed as ‘securing comprehension’.

In Section 4.1.4 *Quotative you know*, it was briefly mentioned that instances where *you know* occurs after a direct speech report are not categorised under ‘quotative *you know*’ if that is not their primary function. There are a couple of such instances in the ELFA corpus that are categorised as *you know* securing comprehension. Although Müller (2005: 174) analyses *you know* occurring after reported speech as “imagine the scene”, in this thesis and in this particular function, *you know* seems to check or ask if the addressee has understood what the speaker is trying to imply with the previous quotation, i.e. the speech act of the utterance:

- (63) <S2> every time we were (xx) were s- told that we know that there are these dictators that development dictators and they are abusing human rights , he shifted er over (glance) and he said well but they increase er lit- the literacy rate *you know* , and this is this is this is a very serious critique of the left and (UDEFD070)

It is debatable if *you know* is a part of the quotation itself or if it marks the ending of the quotation in example (63). It is also debatable whether the speaker uses *you know* to ask the addressee to imagine the scene or to see the implication. However, although *you know* can be considered as an invitation for the addressee to imagine the described scene, it more prominently seems to check that the

addressee understands the implication or the speech act of the utterance – in this case the contradiction between abusing human rights and the increase in literary rate. In fact, checking that the addressee has understood the speaker's implication seems to be a more prominent function especially in argumentative discourse, like in example (63) and in many other instances of *you know* securing comprehension.

In conclusion, when ELF speakers use *you know* to secure comprehension, they do it because they want to check that the addressee is able to understand the speaker or that the addressee is able to understand what the speaker is implying with the previous utterance, i.e. the speech act of the utterance. *You know* securing comprehension occurs primarily during discussions and in argumentative discourse. In addition, *you know* usually occurs at the end of a clause functioning like a tag question. In the majority of cases, *you know* does not receive a verbal response, but as example (61) demonstrates, there is a possibility that the response the speaker receives may be non-verbal. In cases where *you know* does receive a verbal response, it often leads to the speaker yielding the turn to the next speaker. However, the transcripts of the ELFA corpus are not always enough to define if the next turn is selected by the current speaker or the new speaker. Nevertheless, in all instances of *you know* securing comprehension, *you know* is used to initiate the addressee's involvement, which is why it functions on interactional level.

4.1.8 Appeal for understanding

'Appeal for understanding' and 'securing comprehension' share the same feature that the addressee is invited to understand the speaker's meaning because s/he does not explicitly do so. Müller (2005: 181) explains that the difference between the two is that in 'appeal for understanding', "*you know* appeals to the hearer on behalf of the speaker rather than on behalf of the narrative", whereas in 'securing comprehension' the situation is the opposite. The other aspect that sets the two categories

apart is the speaker's level of certainty. Müller (2005: 181) points out that the category 'appeal for understanding' is connected to 'lexical or content search' as 'appeal for understanding' often deals with searching for a correct word or content as well. However, in 'appeal for understanding', the speaker cannot or does not try to search for the missing piece of the utterance but appeals to the addressee to understand nonetheless (Müller 2005: 181–182). In other words, 'appeal for understanding' is a category where the function of *you know* is to ask the addressee to fill in the missing pieces because the speaker is for one reason or another unable to do so.

In the ELFA corpus, all the instances of *you know* occur during discussions, and it is the least frequent function of *you know* with seven instances. All instances of *you know* occur in context where the speaker is having trouble being coherent or finding proper words or expressions. Twice, the speaker explicitly voices his/her uncertainty, like in example (64):

- (64) <S6> yeah , er what er what which transaminase </S6>
 <S5> er E-S-T and E-L-T <S6> yes </S6> erm , <S6> and [er] </S6> [they] tell you liver dimension they show liver dimension in this case it's er just we had just the er the E- E-L-T er that was elevated not the (xx) i don't know mhm the other one , *you know* , i don't know </S5> <P:05>
 <S2> i think it was interesting it's er here in the book er that er if er ALAT is less than 200 (USEMD080)

Example (64) is from a seminar discussion and in the excerpt, speaker S5 clearly states his difficulties in finding the proper word, *I don't know mhm the other one*, and *you know* appeals to the addressee for understanding despite the lack of words. *You know* might even be regarded as an appeal for the others to step in and help finding the proper word. However, the two long pauses surrounding *you know* and the repeated *I don't know* after *you know* do not encourage S6 or anyone else from the seven participants to verbally fill in the missing piece.

Example (64) is similar to example (11) from Müller (2005: 181), where the speaker is searching for the English word for *Kopftuch*. In example (64), speaker S5 is explicitly letting the addressee know that he is unable to remember the term he needs: *I don't know mhm the other one*.

Thus, he appeals to the addressee for help, which makes *you know* function on interactional level rather than on textual level. Moreover, *you know* is preceded and especially followed by a pause that lasts for two seconds or more, which is a long enough pause for the addressee to initiate a turn and actively help the speaker to find the term he is searching (cf. Müller (2005: 181). Why speaker S5 does not receive help, however, remains unclear; maybe the other participants do not know the correct term either, maybe they know what S5 is referring to despite the lack of words or maybe it is not so important that the others feel the need to say it aloud.

There are four instances of *you know* marking an appeal for understanding where it occurs before or after a conjunction. In these cases, the speaker is giving a proposition in the form of *X or Y* or *X so Y*, where Y is left unclear. In example (65), the second *you know* demonstrates such a case:

- (65) <S1> i i'm sure it <S6> [mhm] </S6> [helps] the child to learn about perspectives and how <S6> [mhm] </S6> [to put] things together and how things are three dimensional <S6> [mhm-hm] </S6> [instead] of you know just pictures <S6> mhm </S6> pictures in books or *you know* <S6> mhm-hm </S6> but er <SS> @@ yeah mhm-hm </SS> <P:06> but then er mhm children have two parents you know <S6> [mhm] </S6> [so] i think what the the idea behind the you know the feminist movement (USEMD060)

In example (65), speaker S1 argues for children's construction toys and compares them to 2D objects. The comparing proposition consists of *X or Y*, where X is *pictures in books* and Y is replaced by *you know*. S1 attempts to continue with *but er*, but leaves it unfinished and therefore leaves it to the addressee to understand his meaning. Perhaps it is S6's response *mhm-hm* prompted by *you know* that encourages S1 to carry on with the next part of the argument.

The two remaining instances of *you know* occur after the verbs *BE* and *HAVE*. In both instances, the speaker is having difficulties finding the proper words or content before *you know*. In example (66), *you know* precedes the verb *BE*:

- (66) <S1> yeah the thing is er i think erm it needs to erm to develop some kind of erm er some kind of point for for this website so that it the the audience that are going to that so that they feel that this is so they it's not just erm for for people who are <NS5> yeah </NS5> *you know* </S1> <NS5> it might take a lot of people off as well (ULECD020)

As the excerpt shows, speaker S1 is having difficulties being coherent and does not explain to what kind of *people* she is referring. Perhaps NS5's *yeah* implies S1 that the addressee has understood her meaning despite the fumbling and uses *you know* to appeal for understanding.

In conclusion, when *you know* marks appeal for understanding, the speaker is for one reason or another unable to convey his/her meaning and uses *you know* instead to ask the addressee to understand the speaker's vagueness. Although similar to 'securing comprehension', 'appeal for understanding' is also connected to 'lexical or content search' because in these categories the speaker is having trouble finding the proper words or content or being coherent. All the instances of *you know* occur in the ELFA corpus during discussions, and it is the least frequent function of *you know* in the data.

4.1.9 Reference to shared knowledge

There are 10 instances of *you know* marking reference to shared knowledge, which makes it one of the least frequent function in the ELFA corpus. Most instances occur during discussions. To know that the speaker refers to shared knowledge in the ELFA corpus requires some inference of the speaker's thoughts: what the speaker knows or assumes the addressees to know of the subject at hand. Due to the academic nature of the utilised data, it is reasonable to infer that the speaker does not have a close relationship with the addressees nor that s/he shares many past events or experiences with them (cf. Erman 2001: 1346). However, it is equally reasonable to assume that all participants have basic knowledge of the subject at hand – either because the subject should be generally known in the academic circles or because it has been introduced at some point during the present or previous

lecture/discussion session/etc. Thus, similar to Müller (2005: 178), it can be argued that in the ELFA corpus the speaker knows that the addressees have or should have “the relevant information”.

The overall function of *you know* in this category is to ensure that the addressee is able to access the relevant knowledge that they (should) share, and thus understand the point the speaker wants to make (cf. Müller 2005: 178). In most cases, *you know* refers to academic knowledge that is or should be known to everyone, and occasionally the speaker explicitly reminds the addressee(s) that what s/he is about to say next has been discussed before:

- (67) <S1> [<COUGH> i just say another thing and then you] can answer and also this research that you refer to which stated that these er civic values are permanent <S2> [yeah] </S2> [you] know that they are transmitted basically through generations this is also somehow v-very s- well i think er q- questionable finding because (ULECD140)

Example (67) is from a lecture discussion, where speaker S2 has just given a presentation on his PhD study and hence, the subject should be known to everyone. Therefore, the purpose of *you know* is to remind S2 as well as the other addressees that they all know what it means that *civic values are permanent*, which is *that they are transmitted basically through generations*. Being able to remember that relevant information is important in order to understand the point S2 wants to make, i.e. the critical feedback.

Example (68) is from the beginning of a lecture. This time the shared knowledge is not about the subject of the lecture but about the academic world:

- (68) <S1> my name is <NAME S1> , and i'm going to give you a short lecture on finnish economic and social history , and i'm sorry about the last thursday i was supposed to be here then but i wasn't and , in a way it was not my fault because i had too many duties in the university administration and *you know* the administration always , goes before , education , so that sorry for that and and how many student there's <S2> [(xx)] </S2> [supposed] to be </S1> (ULEC020)

In example (68), the speaker is apologising for his absence from a previous lecture but, on the other hand, he justifies it by reminding everyone of the general truth that in the academic world,

administrative responsibilities come before education. Whether the addressees are aware of this general truth is irrelevant: if they did not know it yet, they know it now.

In Müller's data (2005: 180), few non-native speakers respond verbally to *you know* marking reference to shared knowledge, whereas with native speakers, it is rare that *you know* does not receive a response. The same non-responding tendency is present in the ELFA corpus as there are only two instances of *you know* that receive a response. Example (69) is one of them:

- (69) <S8> that was quite interesting to hear that the first law was 19 er 1889 <S4> mhm-hm
</S4> *you know* that er that gentleman <NAME> told us or show show us some slides <S2>
uh-huh </S2> powerpoint slides </S8> <S2> wh- what kind of law was it then </S2>
(USEMD160)

Speaker S8 gives a proposition before *you know* and the utterance after it functions as a reminder to a past event that ought to be shared by everyone. S2's minimal response *uh-huh* conveys S8 that he is able to remember the occasion.

Although most instances of *you know* relate one way or another to the academic world, once an ELF speaker seems to refer to general knowledge:

- (70) <S1> there's this sort of metaphor from henry henry miller actually this er standing still like
a humming bird like you *you know* a humming bird can sort of er stand still in the air by
doing an enormous amount of work the wings are flapping all the time and so if er there
there's a huge huge amount of work (USEMD130)

In example (70), speaker S1 uses the metaphor *stand still like a hummingbird* by an American writer Henry Miller. The purpose of *you know* is to introduce a reminder to general knowledge of how hummingbirds fly, because remembering it is important in order to understand what S1 means by the metaphor.

To summarise, *you know* marking reference to shared knowledge is used by the speaker when s/he wishes to ensure that the addressee is able to access the relevant knowledge they both share (or at least should share), and thus is able to understand the point the speaker wants to make. Based on

the nature of the ELFA corpus, the shared knowledge is mostly academic, either about the subject or the academic world in general. Most instances of *you know* in this category occur during discussions but rarely *you know* receives a confirming verbal response from the addressee.

4.1.10 Acknowledge that the speaker is right

There are 26 instances of *you know* marking ‘acknowledge that the speaker is right’. With this function, the speaker uses *you know* to ask the addressee to acknowledge that the proposition the speaker gives before *you know* is correct, or to state that the addressee is right or that at least s/he has the right for the opinion (see also Müller 2005: 183–186). In the ELFA corpus, *you know* occurs in this function only during discussions, and since it comes after the speaker’s proposition, it occurs in clause-final position.

Example (71) is from a doctoral defence discussion, where speaker S2 acts as S1’s opponent. In the excerpt, S2 first asks a question about a quotation S1 has not included in his dissertation. Although S1 begins to answer, S2 takes the floor by adding his opinion about the excluded sentence:

- (71) <S2> the great german er marxist philosopher ernst bloch once said he said <READING ALOUD> everything that is in the superstructure is al- also in the base except for the base </READING ALOUD> <S1> @@ </S1> okay , er , why did er you cut out that sentence i quoted about giddens </S2> <S1> er you know er [(xx)] </S1> <S2> [i think] it was actually very good *you [know]* </S2> <S1> [it was] er well er , i sent out my copy out for copy (xx) among colleagues and , i found out that people were coming hard on me on that aspect (UDEFD070)

The first *you know* spoken by S1 marks content search, whereas the second *you know* spoken by S2 marks acknowledgement that the speaker is right where S2 is either asking S1 to agree with him or stating that he is right, ergo S1 should agree with him.

Similar to other categories where *you know* functions on interactional level, most of the instances of *you know* do not receive a response from the addressee. For example in example (71),

S1 takes the next turn during S2's *you know*, and therefore one cannot be certain if it is *you know* or the whole utterance that prompts S1 to answer. However, example (72) demonstrates a clear case where the speaker receives a back-channel as a response for his argument:

- (72) <S11> yeah i would like to comment to the last point that okay people have access but do they have any special use for these it's always a question what what kind of uses they have for actually in rural areas for mobile phones it's it's not something to be taken for granted that <S34> it's [yeah] </S34> [you] have access you have a need </S11>
 <S34> you make business through mobile phones you <S11> [yeah yeah yeah] </S11> [call your neighbour] *you know* what you can <S11> [yeah yeah] </S11> [(xx)] (xx) </S34>
 (CDIS08B)

For example (72), the ELFA recordings were listened to to make sure that *you know* is about acknowledgement and not about, e.g., linking proposition together. S34 responds to S11's question by giving examples of what people in rural areas do with mobile phones. *You know* is said with an even pitch after the examples, which indicates that the speaker is sure about being right. Moreover, S11 response *yeah yeah* soon after *you know*, which implies that he does indeed agree with S34.

There is also one instance of *you know* in the ELFA corpus where it occurs inside a quotation but functions as a request or statement for the addressee to acknowledge that the speaker is right (see also Müller (2005: 185); 4.1.7 *Securing comprehension*). In example (73), there are in fact four instances of *you know* and none of them share the same function. Furthermore, three of them occur inside a direct speech report. The second *you know* functions as 'acknowledge that the speaker is right':

- (73) <S1> whenever i have discussions with my friends about for example what russia is you know some of my friends say well there is a place in ukraine there is a stone which says this is where russia started so <S2> [@@@] </S2> [don't argue anything] else *you know* how can you say that you know these things are constructed if there is even a place which says that this is you know this is where the russian empire started from so this is also the kind of pain that i have been facing because when you have realised these things and you are trying to pro- show your point then you are faced with this kind of a <S2> mhm </S2> erm argument (USEMD280)

As Erman (1987: 121) points out, *you know* may have more than one function, which seems to be the case in example (73). The first *you know* is categorised as ‘linking propositions together’, the second as ‘acknowledge that the speaker is right’, the third as ‘highlighting certain elements’ and the fourth as ‘marking repair’. The second *you know* comes after a reported proposition *so don’t argue anything else*, and the function of *you know* is to emphasise the proposition as well as the reported speaker’s certainty that s/he is right.

In example (74), the speaker is either asking for the addressee’s agreement or stating this is his opinion and the addressee should believe him to be right:

- (74) <S2> and i think the interlinkage the configuration of these two questions what is the state and what is an NGO in the question of north south it's very important <S1> mhm </S1> and i think you should look into writing <S1> [i will] </S1> [it that] article *you know* , because i think it's incredibly interesting (UDEFD070)

Speaker S2 gives a proposition *I think you should look into writing [it that] article*, and although S1 gives a confirmative response before S2 has finished the proposition, it seems that S2 wishes to emphasise him being right just to be on the safe side.

To conclude, when *you know* function as ‘acknowledge that the speaker is right’, its purpose is to ask the addressee to acknowledge that the proposition the speaker gives is correct, or to state that the addressee is right or at least has the right for the opinion. In the ELFA corpus, *you know* occurs only during discussions and at the end of a clause in this function. Moreover, although *you know* invites the addressee to take part in the conversation, it rarely receives a verbal response from the addressee.

4.1.11 Unclassified instances

There are 17 instances of *you know* that fall under the category ‘unclassified instances’. Müller (2005: 188) notes that in such cases where the speech is unintelligible or the amount of context is insufficient, it is not possible to determine the function of *you know* as a discourse marker. Similarly, unclassified instances in the ELFA corpus consist of cases where the uttered *you know* is either unclear, which is marked in the transcript as *(you know)*, or there is not enough context to determine the function of *you know*.

As mentioned before, there are 11 unclear instances of *(you know)*, which are not analysed nor divided into discourse and non-discourse markers. The remaining seven instances, however, are discourse markers, but their function is difficult to determine with certainty due to the lack of context. With six of those instances, *you know* is preceded or followed by a stretch of talk that is unintelligible, marked in the transcript as *(xx)*, and the length of *(xx)* may vary from one word to multiple clauses in the ELFA corpus. There is also one instance of *you know* where its lack of classification is due to lengthy overlapping talk by several other ELF speakers, and the speaker who uses *you know* stops before the argument is made. This instance is shown in example (75), where the latter *you know* is left unclassified:

- (75) <SS> [(xx)] </SS>
 <S3> [i i'm i'm always curious if they do that you know because it is no use if you just talk about it among amongst yourselves because then nothing else *you know*] </S3>
 <S8> he said something that they were just criticising the methods and things [like that] </S8> (USEMD200)

In example (75), speaker S3's entire turn is spoken simultaneously with other speakers (<SS>) and the proposition S3 makes *because then nothing else* is left unfinished. Perhaps she was searching for a word or content, or perhaps she wanted to highlight the point she was making but gave up her turn due to the simultaneous talk. Nevertheless, because the proposition is left unfinished and the speaker

does not continue the topic during her next turn, one cannot say for certain what the function of *you know* is in example (75).

4.2 Comparison to previous literature

The first research question was answered in the previous section by analysing all the instances of *you know* that function as a discourse marker in the ELFA corpus, and categorising them into different functions. The second research question, in turn, is answered in this section by comparing the results of the analysis to the previous studies of non-native English speakers' use of *you know* that are important for this study, i.e. Müller (2005) and House's (2009) studies. In addition, as some previous studies on native English speakers' use of *you know*, particularly Erman's (2001), were also used as a reference point in the previous section, some comparison is also made to those studies as well.

The first function of *you know* on textual level is *you know* marking lexical or content search. The results of the analysis of the ELFA corpus mostly support Müller (2005) and House's (2009) studies. When analysing lexical or content search, Müller (2005: 158–160) concentrates on pauses and argues that the position of a pause does not determine the target of the search (cf. Östman 1981). In the analysis of the ELFA corpus, the position of the pause – if filled pauses are regarded as one – was not the primary indicator for the kind of search the speaker was doing: both search types include cases where a pause follows or precedes *you know*, as well as cases where there was a pause before and after *you know*. Thus, the result supports Müller's (2005) observation that filled or unfilled pauses do not determine the type of search the speaker is doing. In addition, there were many instances where zero pause occurred near *you know* in the ELFA corpus. However, this can be due to the transcripts of the ELFA corpus where pauses lasting less than two seconds are not marked (ELFA Transcription Guide 2004: 1).

Most instances of *you know* were relatively apparent cases of either lexical or content search in the ELFA corpus. However, there were some instances where clear distinction was not possible to make. This also supports Müller's (2005: 160) argument that the category 'lexical or content search' is a continuum that consists of clear instances of both search types, as well as instances that have elements of both.

House (2009: 186), on the other hand, argues that in cases where the speaker is being inconsistent or having difficulties finding the appropriate word or formulation, "*you know* occurs in mid utterance and also inside nominal, verbal and adverbial groups". In addition, House (2009: 187) argues that *you know* also functions as a focus marker when the speaker is fumbling for words or content. The results of the ELFA corpus partly support House's (2009) argument as *you know* occurs primarily within an NP or a VP, and often in mid-utterance when the speaker is doing lexical search. Moreover, *you know* also functions as a focus device in several instances when the speaker is being incoherent. However, when *you know* marks content search, it usually occurs in clause-initial position, which argues against House's (2009) observation. In fact, the results of the ELFA corpus are somewhat similar to Erman (2001), who studies native speakers of English. Erman (2001: 1344) observes that as a marker of lexical search, *you know* occurs within a phrase after a determiner, whereas as a marker of content search, *you know* usually occurs at the beginning of a clause. However, Erman (2001: 1344) defines *you know* to usually occur after a conjunct or disjunct (conjunctions included), but in the ELFA corpus, *you know* occurs many a time before conjunctions. In fact, except for conjunctions, there were only a few instances where *you know* occurred before or even after any other conjuncts or disjuncts in the ELFA corpus.

You know marking repair is the second functions of *you know* on textual level. The results of the ELFA corpus support Müller's (2005: 160–162) findings that *you know* is a marker of repair, which consist of cases where a complete or unfinished word is substituted with another after *you know*, and of cases where nothing is repaired but only repeated. However, the results of the ELFA

corpus also show that because Müller (2005) does not take into account cases where the syntactic structure is changed after *you know* (cf. Erman 2001: 1345), her definition of repair is inadequate at least for analysing the ELFA corpus. As for House (2009: 186), the results of the analysis support her notion that *you know* marking repair occurs in mid-utterance and inside verbal and nominal groups. However, there are only a few instances of *you know* occurring after AdvPs or PrepPs that function as adverbials in the ELFA corpus. Thus, it can be argued that *you know* marking repair does not usually occur inside adverbial groups in the ELFA corpus (cf. House 2009: 186).

There are some instances of *you know* in the ELFA corpus where the distinction between ‘marking lexical or content search’ and ‘marking repair’ is not easily made (see example (32)). These cases support House’s (2009) manner of not differentiating lexical or content search and repair from one another. Nevertheless, many other examples from the ELFA corpus show that such a distinction can be made. In fact, Erman (1987: 121) and Müller (2005: 174) also note that although the functions of *you know* can be labelled and categorised, it does not mean that *you know* only has one set of function at a time, but it can have elements of other functions as well. Thus, it is the most salient function of *you know* that determines to which category *you know* belongs (Müller 2005: 174).

‘Introducing an explanation’ is the third category that functions on textual level. The results of the analysis of the ELFA corpus support Müller’s (2005: 167) findings that as an introduction to explanation, “*you know* is used to indicate ideational relationships between statements or concepts which precede it and those which follow it”. Therefore, as Müller (2005: 167) points out, *you know* also functions on textual level. The form of explanation, though, is different in the ELFA corpus compared to Müller’s (2005) study. In Müller’s (2005: 166) data, the majority of instances are explanations where “the speaker mentions something, a concept or an idea, or gives his/her opinion, and then decides that s/he has to express it in different (and perhaps more) words to make it plain what s/he meant.” These cases are similar to the notion of rephrasing previous discourse used in this thesis (see 4.1.3 *Introducing an explanation*). However, although rephrasing is not an uncommon

form of explanation, it is nevertheless the least common in the ELFA corpus. On the other hand, the most common forms of explanation in the ELFA corpus, modification and parenthetic comments, are rare in Müller (2005): modification is not (at least explicitly) present in her forms of explanation and the presence of parenthetic comments in this particular category is also debatable in her study (see the discussion in 2.2.3 *Introducing an explanation*).

Another interesting difference between this study and Müller's (2005) is that *you know* introducing an explanation is the most frequent function in Müller (2005), whereas it is not that common in this study. The reason for this may be due to the additional category 'linking propositions together' that is not present in Müller (2005), because 'linking propositions together' also fits into Müller's (2005: 167) definition of *you know* introducing an explanation: that the function of *you know* introducing an explanation is to point out "ideational relationships between statements or concepts which precede it and those which follow it." The difference, however, lies in plainness: an explanation is introduced when the speaker is afraid of being unclear, whereas propositions are linked when the speaker is being clear but wants to emphasise the connection between the propositions.

In regards to House's study (2009: 184), the results of the ELFA corpus support her views that *you know* introducing an explanation may act instead of conjunctions to underline and make more explicit the ideational relations between the utterances adjacent *you know*. However, whereas House (2009) emphasises the relation between *you know* and the conjunctions *and*, *but* and *because*, the conjunctions relevant for this function seem to be *or* and *because* in the ELFA corpus. When *you know* introduces a modification or a rephrased utterance, it often acts instead of *or* (e.g. examples (35) and (36)). Sometimes, if the syntactic structure of the previous utterance allows it, *you know* introducing a parenthetic comment could be seen as replacing the conjunction *because* (e.g. example (37)). In the rest of the instances, however, *you know* seemed to act on its own right and not in conjunctions' stead, marking an additional clause that is ideationally related but not always syntactically appropriate for the utterance preceding *you know*.

‘Quotative *you know*’ is the fourth category of *you know* functioning on textual level, and it is a function discussed in Müller (2005) but not in House (2009). The approach to ‘quotative *you know*’ in this thesis slightly differs from Müller’s (2005). Müller (2005: 168–170, 174) defines *you know* functioning as an enquoting device when it occurs before a quotation or in between quotations, but excludes instances where *you know* occurs after quotation and analyses them as “imagine the scene” (‘securing comprehension’ in this thesis). Müller (2005: 174) explains that although *you know* could be viewed as ending a quotation, it seems to have another, more salient function, i.e. asking the addressee to imagine the scene. The same reasoning is utilised in this thesis as well, but it consists of instances of *you know* in between and after quotations. In other words, when *you know* occurs within or after a quotation, it is not analysed as ‘quotative *you know*’ because it seems to have another, more salient function than introducing a continued or finished quotation (see, e.g., example (63) in 4.1.7 *Securing comprehension* or example (73) in 4.1.10 *Acknowledge that the speaker is right*).

However, ‘quotative *you know*’ being one of the least common functions in the ELFA corpus is similar to Müller’s (2005) results. In fact, only Americans used *you know* to introduce reported speech in Müller’s (2005) data. What is also interesting is that in all the examples Müller (2005: 169–170) gives, *you know* co-occurs in close vicinity with the structure *BE + like*, where *BE* represents all the various forms of the verb. However, this structure occurred only once in the ELFA corpus (see example (41)), whereas in most cases *you know* co-occurred with the explicit quotative verb *SAY*. This lack of use of *you know* with the structure *BE + like* when introducing reported speech may indicate that the structure is not explicit enough for ELF speakers to introduce a speech report, but a more distinct quotative verb is needed (for further analysis on *like* as a marker of reported speech, see, e.g. Romaine and Lange 1991).

You know linking propositions together is the fifth function on textual level, and it is a function mentioned in House (2009) but not in Müller (2005). The results of this thesis mostly correspond to House’s (2009) results. House (2009: 184) argues that *you know* may function “on its own as a

relational phrase” to make the relation between clauses more explicit without the presence of conjunctions, *and*, *but* and *because* in particular. In the ELFA corpus, *you know* often links together an argument and the utterance that supports the argument said after *you know*. Thus, in these cases, *you know* could many a time be replaced by *because* that would serve the same purpose, i.e. linking together – in Erman’s (2001: 1343) words – “the speaker’s position and the backing up of it.” However, based on the instances in the ELFA corpus, *you know* could also be seen as replacing or functioning in a similar manner as other cohesive devices, such as linking words or even punctuation marks in written text. Thus, the results of this thesis support House’s (2009: 184) notion that *you know* may function “on its own ... to make implicit coherence relations more explicit”. However, they also suggest that *you know* can be seen to have similar functions to many other cohesive devices just as easily as to conjunctions. In other words, the emphasis should be on the link *you know* creates, not on certain conjunctions *you know* seems to replace.

‘Highlighting certain elements’ is the sixth and last category to function primarily on textual level in the ELFA corpus. As Müller (2005) does not mention *you know* having a highlighting function, the results of the ELFA corpus are not comparable to her study. It is, nevertheless, interesting that the most frequent function of *you know* in this study is absent in Müller’s (2005). This, however, may be due to the different natures of the used corpora: the GLBCC corpus Müller (2005) utilises consists mostly of narratives and light conversation, whereas the ELFA corpus consists of academic discourse where sharing academic knowledge and argumentation is commonplace, and therefore, highlighting important elements in discourse is important.

House (2009), on the other hand, credits *you know* having a highlighting function. However, House (2009: 181) argues that *you know* primarily highlights conjunctions co-occurring with it and makes more salient the relations these conjunctions create. It is clear that in the ELFA corpus *you know* also highlights the co-occurring conjunctions, yet the approach in this study to conjunctions is

different from House (2009). The excerpt below is from House (2001: 181, original emphases) and it serves as an example to demonstrate the difference:

- (76) N: But no **but** look **you know** I agree **but** on the other side look at all the different nationalities that live in Germany and the all speak German and proper English too **but** all have their little niches **you know** they all have their culture like the Turks and the and the what ever

As previously mentioned, House (2009: 181) argues that the function of *you know* is to highlight and make salient the relations established by conjunctions co-occurring with *you know*. Thus, the function of *you know* in example (76) is to highlight the contrastive function of the conjunction *but*. Moreover, based on the last *but* and the latter *you know* in example (76) from House (2009: 181), the stretch of talk between *but* and *you know* may be as long as a clause. However, had example (76) been from the ELFA corpus, the first *you know* would have been defined as *you know* highlighting a conjunction, whereas the second *you know* would have been defined as linking propositions together. This is because it occurs between two largely independent propositions, *but all have their little niches* and *they all have their culture like the Turks and the and the what ever*. Hence, the difference between these two studies is that in this study the mere presence of a conjunction is not enough to consider *you know* as a highlighter, but the conjunction must occur in close vicinity, preferably immediately or almost immediately next to it (see, e.g., examples (50)–(52)).

Moreover, in House (2009) and Erman (1987, 2001), *you know* highlighting certain elements is also connected to Halliday's (1985) *theme*. Erman (1987: 130–131) explains that when *you know* occurs after fronted adverbial, it organises the thematic structure of the sentence and introduces the upcoming argument. In a later study, Erman (2001: 1343) refers to *you know* occurring in the same situation as *you know* “marking, possibly also highlighting, the fronted adverbial, which functions as a scene setter for the ensuing course of events.” In addition, Erman (2001: 1342) explains that “the speaker urging the listener to accept part of the information as ‘known’ or ‘given’” is one example of

you know marking, i.e. highlighting, “certain elements in the thematic structure”. House (2009: 183) also notes that *you know* seems to create a link between Halliday’s (1985) theme, which she paraphrases as “the heading to what I am saying”, and what is said after *you know*. However, based on the ELFA corpus, it seems that equal recognition should be given to the *rheme* or what comes after *you know*, because it holds new or important information of the proposition and hence the point the speaker wants to say. He and Lindsey (1998: 134) make a similar observation, stating that *you know* increases “the status of information in terms of saliency, importance, and newness, thereby orienting themselves and their interlocutors to information that enjoys that status, which in turn structures interactional participation in specific ways.” There are two important aspects in the quotation by He and Lindsey (1998: 134). First, it points out that the rheme in the utterance is just as important as the theme, and second, it shows that highlighting also invites the addressee to participate, i.e. it functions to some extent on interactional level. In other words, although *you know* highlighting certain elements is categorised as functioning primarily on textual level, it has interactional elements as well.

In addition, Erman (2001: 1345) notes that *you know* can function at the same time as a turn-taking device and a highlighter. According to Erman (2001: 1345–1346), this function is particularly common with native English adolescent speakers and in most cases the selection for the next speaker is done by the speaker him/herself. Turn-taking function is not mentioned in Müller (2005) or House’s (2009) studies on non-native speakers of English, but it is present in the ELFA corpus. Sometimes the speaker is selected by the previous speaker but most of the time it is done by the speaker him/herself. In all instances, the function of *you know* is also to highlight the point the speaker wants to make.

You know securing comprehension is the first category that functions mostly on interactional level in the ELFA corpus. In this thesis, ‘securing comprehension’ corresponds to two different categories in Müller’s (2005) study: “imagine the scene” and “see the implication”. The results of the ELFA corpus support Müller’s (2005: 171) argument that when *you know* functions on interactional

level, it is used to elicit some kind of involvement or co-operation from the addressee. However, it also seems that in the ELFA corpus, “see the implication” is a more prominent feature than “imagine the scene” (e.g. example (63)). This may be once again due to the different natures of the used corpora: the data Müller (2005) utilises is based on silent movie narratives, whereas the ELFA corpus consists of academic discourse. Thus, as argumentation is a common mode in academic discourse, it is natural that the ELF speakers in the ELFA corpus use *you know* to secure that the addressee comprehends what they are saying or implying.

Contrary to Müller (2005), House (2009) does not support the notion of *you know* eliciting involvement, co-operation or any other mode of intersubjectivity in non-native English discourse. Therefore, none of the functions operating on interactional level are supported by House (2009). Consequently, House (2009: 188) argues that ELF speakers do not receive or even expect to receive a response when using *you know*. However, as the analysis of *you know* in the ELFA corpus showed, *you know* securing comprehension elicits at least mental involvement from the addressee and sometimes it also receives a response that varies from back-channelling to turn-taking.

The second category functioning on interactional level is ‘appeal for understanding’. The findings in this thesis supports Müller’s (2005) definition of *you know* marking appeal for understanding when the speaker cannot find the right words or content but asks the addressees to understand what s/he is trying to convey nevertheless. Moreover, Müller (2005: 182–183) observes that sometimes non-native speakers use *you know* in this sense with “an apologetic note” to appeal the addressees’ sympathy, but as intonation contour is not marked in the ELFA transcripts, the speaker’s voice quality is not taken into account in this thesis. It is interesting, though, how few of the instances of *you know* function as an appeal for understanding in the ELFA corpus. Perhaps this is again due to the academic nature of the ELFA corpus, the speakers trying to be precise with their delivery.

‘Reference to shared knowledge’ is the third category that functions on interactional level in the ELFA corpus. The analysis of *you know* supports Müller’s (2005: 178) observation that the function of *you know* in this category is to ensure the addressee’s ability to remember certain information and, therefore, understand the point of the speaker’s proposition. It also follows the line of *you know* not often receiving a verbal response from the addressee that was present in Müller’s (2005: 180) data. This, in turn, gives some support to House’s (2009: 188) argument that ELF speakers and addressees do not consider *you know* as an invitation to get involved in the conversation by back-channelling or replying to *you know*. However, although ELF addressees “are NEVER specifically addressed” in House’s (2009: 188) data, it cannot be generalised to cover all ELF speakers as *you know* occasionally receives a response in the ELFA corpus (see also 4.1.7 *Securing comprehension*).

‘Acknowledge that the speaker is right’ is the fourth and final function of *you know* on interactional level. The analysis of the ELFA corpus conforms to Müller’s (2005: 185) observation that *you know* rarely receives a response when it functions as a marker of ‘acknowledge that the speaker is right’. Hence, similar to other categories of *you know* functioning on interactional level, it also supports House’s (2009: 188) argument that *you know* does not seem to function as an invitation to get the addressee involved in the conversation. However, although once again the ELF speakers do not often receive a response in the ELFA corpus, the times they do receive one oppose House’s (2009: 188) findings on ELF speakers who never seemed to be “specifically addressed”.

In conclusion, the functions of *you know* in the ELFA corpus mostly seem to agree with Müller (2005) and House’s (2009) studies. However, there are some interesting differences as well. Compared to Müller (2005), this study also divides *you know* into ten different functional categories, but it is interesting that the most frequent function in Müller (2005) is not at all common in the ELFA corpus. It is also interesting that the most frequent function in the ELFA corpus is not part of the functions set by Müller (2005).

Another interesting point is that based on the quantitative data in Müller (2005: 190; figures 4.1.a and 4.1b), non-native speakers seem to use *you know* equally on both textual and interactional levels. However, the division is not as equal in the ELFA corpus. The vast majority of the instances of *you know* in the ELFA corpus function on textual level with 87.6%, leaving the remaining 12.4% of the instances of *you know* to function on interactional level. Furthermore, most of the functions of *you know* that operate on interactional level are the less frequent ones in the ELFA corpus. Figure 2 shows the frequencies of *you know* from most frequent to least frequent in the ELFA corpus:

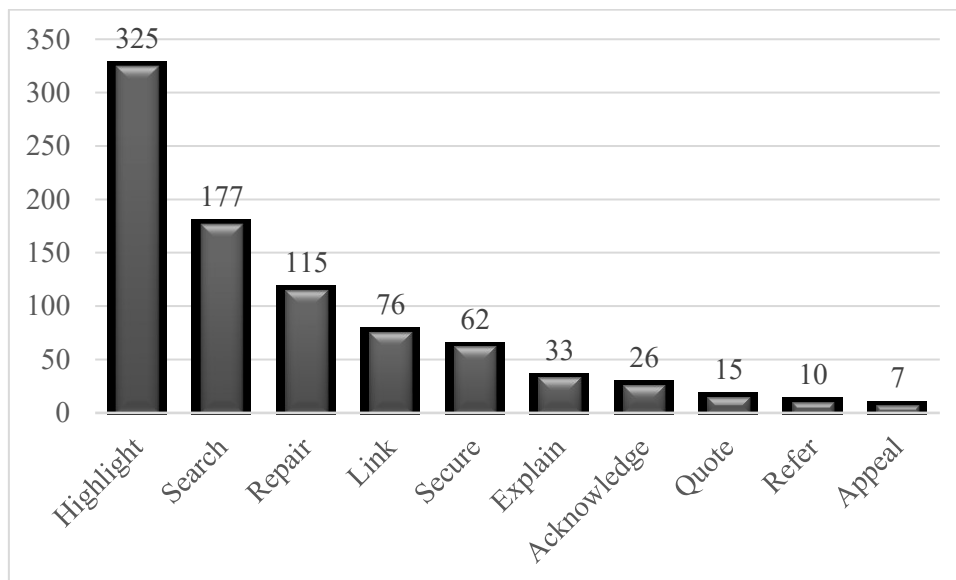


Figure 2. Frequency of *you know* in different functions

The first four categories in figure 2, ‘highlighting certain elements’, ‘marking lexical or content search’, ‘marking repair’ and ‘linking propositions together’, operate on textual level, which means that the fifth frequent category ‘securing comprehension’ is the first category that operates on interactional level. ‘Introducing an explanation’ and ‘acknowledge that the speaker is right’ occur almost as frequently but function on different levels, textual level and interactional level, respectively. ‘Quotative *you know*’ is the only category operating on textual level that is one of the least frequent functions in the ELFA corpus, and yet it is almost as frequent as the last interactional categories, ‘reference to shared knowledge’ and ‘appeal for understanding’, combined.

You know operating primarily on textual level, in turn, supports House's (2009) argument that ELF speakers use *you know* in a self-centred manner, i.e. they focus on the textual aspects of discourse rather than creating a feeling of intersubjectivity with the addressee. However, it must be taken into account that the most frequent function of *you know* in the ELFA corpus is 'highlighting certain element' where the purpose of *you know* is to help the addressee in interpreting the message by emphasising the important part of the utterance. Therefore, even though *you know* highlighting certain elements functions primarily on textual level, it has interactional elements as well, which in turn does not agree with House's (2009) claim that ELF interaction seems to be "a self-centered affair with speakers leaving hearers the freedom of interpretation and ... the freedom to create assumptions." In fact, it seems to be doing the opposite. Furthermore, in the ELFA corpus, ELF speakers occasionally receive a response from the addressee when *you know* functions on interactional level, whereas in House's (2009) data that never occurs. It is also interesting to note that studies on native English speakers (e.g. Erman 2001, 1987) were necessary when analysing the four most frequent functions in the ELFA corpus. Thus, the results of this study show that although there are many functions of *you know* in ELF interaction that researchers of non-native English can agree on, more research is nevertheless needed in studying how *you know* as a discourse marker is used by fluent ELF speakers.

5. Conclusions

The purpose of this study was to take part in the rising discussion of how the discourse marker *you know* is utilised by fluent ELF speakers in academic settings. Hence, the aim of this thesis was to study 1) how *you know* is used as a discourse marker by academic ELF speakers and 2) compare those findings to the studies from Müller (2005) and House (2009) who also utilise corpus data from non-native English speakers with academic background. The results of this study mostly agree with

Müller (2005) and House's (2009) findings, but naturally there are some differences, which is a common trade in analysing discourse markers and their functions (Schourup 1999: 228).

The data used in this study was from the ELFA corpus that consists of ELF interactions in academic settings (ELFA 2008). The type of interactions in the ELFA corpus are various consisting of monologues (e.g. lectures), dialogues (e.g. doctoral defence discussions) and polylogues (e.g. seminar discussions). Therefore, a corpus-driven approach was utilised in this thesis, where according to Tognini-Bonelli (2001: 84), all examples of the analysed linguistic item must be taken into account without any modifications or limitations. For that reason, every *you know* functioning as a discourse marker and spoken by an ELF speaker from all interaction types were included in the analysis.

The first research question was answered by analysing all the instances of *you know* in the ELFA corpus that were defined as discourse markers and were spoken by ELF speakers. These instances were divided into ten different functions that occur in the ELFA corpus, six functioning on textual level and four on interactional level. The results of the analysis show that as a discourse marker, *you know* functions primarily on textual level in the ELFA corpus, meaning that it is oriented towards the textual aspects of discourse. These aspects are, e.g., creating coherence and helping the speaker to plan the following utterance. Thus, the minority of instances of *you know* function on interactional level, where its purpose is to elicit involvement from the addressee, e.g., by asking him/her to understand what the speaker is implying with his/her words.

The second research question was answered by comparing the results of the analysis to the results of Müller (2005) and House's (2009) studies on non-native speakers of English. In addition, the findings of the four most frequent categories in the ELFA corpus were also compared to Erman's (2001, also 1987) studies in particular. This was due to the observation that these instances of *you know* in the ELFA corpus seemed to correspond to Erman's (2001, 1987) notions on *you know* although her data consists of native English speakers. The functions *you know* seems to have in the ELFA corpus mostly correspond to the functions described in Müller (2005) and House's (2009)

studies. However, there are some significant differences as well. For example, the most frequent function of *you know* in the ELFA corpus, ‘highlighting certain elements’, is not present in Müller’s (2005) functional categories. Furthermore, House (2009) argues that *you know* does not function on interactional level in ELF interaction, whereas there are instances of *you know* in the ELFA corpus that are directed towards the addressee. Although these instances are a minority in the ELFA corpus, it still differs from House’s (2009) data where these instances were nonexistent. Moreover, even though the results of this thesis support House’s (2009: 171, 190) argument that ELF speakers use *you know* primarily “as a self-serving strategy”, they do not agree with her claim that ELF speakers give the addressee “the freedom of interpretation and ... the freedom to create assumptions.” In fact, the purpose of *you know* highlighting certain elements, i.e. the most frequent function of *you know* in the ELFA corpus, does completely the opposite: it emphasises the important part of the utterance and thus helps the addressee to interpret the message. Furthermore, although *you know* highlighting certain elements functions primarily on textual level, it also has interactional elements as well.

In this thesis, the analysis of the discourse marker *you know* in the ELFA corpus is primarily based on the ELFA transcripts. Therefore, it would be interesting to study what kind of intonation contour ELF speakers’ use for *you know* and whether it gives any further information on the functions defined in Section 4.1. It would also be interesting to study if the roles of the speakers affect how *you know* is used in academic discourse: is *you know* used differently between students than between a student and a lecturer?

This thesis showed that although there are similarities between this study and other previous studies (Müller 2005; House 2009), there are some significant differences as well. Thus, more research is needed on *you know* as a discourse marker in fluent ELF interactions, academic settings included. This thesis also showed, like many other studies before, that *you know* is not a filler used in spoken language to fill in a gap when the speaker does not have anything else to say. In fact, it

showed that *you know* has an important function to aid the speaker as well as the addressee to form and interpret a message. In other words, *you know* is, and continues to be, relevant.

Works Cited

- Chandler, Daniel and Munday, Rod, eds. 2016. S.v. speech act. *A Dictionary of Media and Communication*. 2nd ed. Oxford: Oxford University Press. Online version.
- Crystal, David. 1988. "Another look at, well, you know..." *English Today* 13: 47–49.
- ELFA = Corpus of English as a Lingua Franca in Academic Settings, The. 2008. Director: Anna Mauranen. Available from: <http://www.helsinki.fi/englanti/elfa/>. (6 May 2017).
- ELFA Transcription Guide. 2004. "ELFA Transcription Guide (7/2004)". Director: Anna Mauranen. Available from: <http://www.helsinki.fi/elfa/elfacorporus/>. (18 March 2017).
- Erman, Britt. 1987. *Pragmatic Expressions in English: A Study of You know, You see and I mean in Face-to-face Conversation*. Acta Universitatis Stockholmiensis, Stockholm Studies in English LXIV. Doctoral dissertation. Stockholm: Almqvist & Wiksell International.
- Erman, Britt. 1992. "Female and Male Usage of Pragmatic Expressions." *Language Variation and Change* 4, 2: 217–234.
- Erman, Britt. 2001. "Pragmatic markers revisited with a focus on *you know* in adult and adolescent talk." *Journal of Pragmatics* 33: 1337–1359.
- Firth, Alan. 1996. "'The discursive accomplishment of normality' : On 'lingua franca' English and conversation analysis." *Journal of Pragmatics* 26: 237–259
- Fox, Barbara A. and Jasperson, Robert. 1995. "A syntactic exploration of repair in English conversation." In *Alternative Linguistics: Descriptive and Theoretical Modes*, ed. Philip W. Davis, 77–134. Amsterdam: John Benjamins Publishing Company.
- Fraser, Bruce. 1996. "Pragmatic Markers." *Pragmatics* 6, 2: 167–190.
- Fraser, Bruce. 1999. "What are discourse markers?" *Journal of Pragmatics* 31: 931–952.
- Fuller, J. M. 2003. "The influence of speaker roles on discourse marker use." *Journal of Pragmatics* 35, 1: 23–45.
- Fung, Loretta and Carter, Ronald. 2007. "Discourse Markers and Spoken English: Native and Learner Use in Pedagogic." *Applied Linguistics* 28, 3: 410–439.
- Gánem-Gutiérrez, G. A. and Roehr, K. 2011. "Use of L1, metalanguage, and discourse markers: L2 learners' regulation during individual task performance." *International Journal of Applied Linguistics* 21, 3: 297–318.
- Halliday, M.A.K. 1985. *An introduction to functional grammar*. London: Arnold.
- Halliday, M.A.K. and Matthiessen, Christian M.I.M. 2004. *An introduction to functional grammar*. 3rd ed. London: Arnold.

- He, Agnes Weiyun and Lindsey, Brian. 1998. "'You know' as an information status enhancing device: arguments from grammar and interaction." *Functions of Language* 5, 2: 133–155.
- Holmes, Janet. 1986. "Functions of you know in women's and men's speech". *Language in Society* 15: 1–22.
- House, Juliane. 2009. "Subjectivity in English as Lingua Franca discourse: The case of *you know*." *Intercultural Pragmatics* 6, 2: 171–193.
- Jucker, A. H. and Smith, S. W. 1998. "Discourse Markers as Negotiating Strategies." In *Discourse Markers: Description and Theory*, eds. A. H. Jucker and Y. Ziv, 171–201. Pragmatics & Beyond. Amsterdam: John Benjamins Publishing Company.
- Jucker, A. H. and Ziv, Yael. 1998. "Discourse Markers: Introduction." In *Discourse Markers: Description and Theory*, eds. A. H. Jucker and Y. Ziv, 1–12. Pragmatics & Beyond. Amsterdam: John Benjamins Publishing Company.
- Liu, Dilin. 2008. "Linking adverbials: An across-register corpus study and its implications." *International Journal of Corpus Linguistics* 13, 4: 491–518.
- Macaulay, Ronald. 2000. "*You know*, it depends." *Journal of Pragmatics* 34: 749–767.
- McArthur, Tom, ed. 2003a. S.v. adverbial. *Concise Oxford Companion to the English Language*. Oxford: Oxford University Press. Online version.
- McArthur, Tom, ed. 2003b. S.v. determiner. *Concise Oxford Companion to the English Language*. Oxford: Oxford University Press. Online version.
- Müller, Simone. 2005. *Discourse Markers in Native and Non-Native English Discourse*. Amsterdam: John Benjamins Publishing Company.
- Polat, Brittany. 2011. "Investigating acquisition of discourse markers through a developmental learner corpus." *Journal of Pragmatics* 43, 15: 3745–3756.
- Redeker, Gisella. 1990. "Ideational and pragmatic markers of discourse structure." *Journal of Pragmatics* 14, 3: 367–382.
- Romaine, Suzanne and Lange, Deborah. 1991. "The Use of like as a Marker of Reported Speech and Thought: A Case of Grammaticalization in Progress." *American Speech* 66: 3, 227–279.
- Sacks, Harvey, Schegloff, Emanuel A. and Jefferson, Gail. 1974. "A Simplest Systematics for the Organization of Turn-Taking for Conversation." *Language* 50, 4: 696–735.
- Schegloff, Emanuel A. 1988. "Goffman and the Analysis of Conversation." In *Erving Goffman: Exploring the Interaction Order*. Eds. Paul Drew & Anthony Wootton, 9–135. Cambridge: Polity Press.
- Schiffrin, Deborah. 1987. *Discourse markers*. Cambridge: Cambridge University Press.

- Schiffrin, Deborah, Tannen, Deborah and Hamilton, Heidi E. 2015. "Introduction to the First Edition." In *The Handbook of Discourse Analysis*, 2nd ed. Eds. Deborah Tannen et al., 1–7. John Wiley & Sons.
- Schourup, Lawrence, 1999. "Tutorial overview: Discourse markers." *Lingua* 107, 3–4: 227–265.
- Tognini-Bonelli, Elena. 2001. *Corpus Linguistics at Work*. Studies in Corpus Linguistics 6. Amsterdam: John Benjamins Publishing Company.
- Östman, Jan-Ola. 1981. *You know: A Discourse-Functional Approach*. Pragmatics & Beyond. Amsterdam: John Benjamins B. V.

Appendix 1. Abbreviations used in the transcript of the ELFA corpus

All abbreviations as well as their explanations are from ELFA Transcription Guide (2004) or from the event descriptions that appear at the beginning of every text document:

| | |
|---------------------------|--|
| <S1> | Utterance begins |
| </S1> | Utterance ends |
| <S2> | ELFA speaker |
| <NS3> | Native speaker |
| <SU> | Unidentified speaker |
| <SS> | Several simultaneous speakers (usually laughter) |
| (text) | Uncertain transcription |
| (xx) | Unintelligible speech |
| @@ | Laughter |
| @text@ | Spoken laughing |
| , | 2–3 seconds pause while speaking |
| . | 3–4 seconds |
| <P: 05> | 5 seconds or longer, rounded up to the nearest second |
| [text] | Overlapping speech (approximate, shown to the nearest word, words not split by overlap tags) |
| <S1> mhm </S1> | Backchannelling |
| er | /öö/ |
| erm | /((ö)m/ |
| ah | /aa/ |
| NATO, EU etc. | Capital letters: only in acronyms |
| <NAME> | Names of participants |
| <SIC> text </SIC> | Nonsense words |
| T-U-C, V-W | Spelling out a word or acronym etc., as letters |
| <READING> text </READING> | Reading aloud |
| <FOREIGN> text </FOREIGN> | Switching into a foreign language |